

AD \_\_\_\_\_

GRANT NUMBER DAMD17-93-J-3003

TITLE: Wavelet Representation for Digital Mammography

PRINCIPAL INVESTIGATOR: Andrew F. Laine, D.Sc.

CONTRACTING ORGANIZATION: University of Florida  
Gainesville, Florida 32611-5500

REPORT DATE: December 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander  
U.S. Army Medical Research and Materiel Command  
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;  
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

THIS QUALITY INSPECTED 8

19970711 125

## FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

Where copyrighted material is quoted, permission has been obtained to use such material.

Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

✓ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Audrey Laine 3-7-97  
PI - Signature Date

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Overview of Contents . . . . .	12
1.2	Publications . . . . .	14
1.3	Responses to Technical and Contractual Issues . . . . .	15
1.4	Study in Progress: An ROC Comparison Between Digital Mammography and Screen-Film Using an Anthropomorphic Breast Phantom . . . . .	16
1.4.1	Introduction . . . . .	16
1.4.2	Method . . . . .	16
<b>2</b>	<b>Body</b>	<b>20</b>
2.1	Generalizing the Discrete Dyadic Wavelet Transform . . . . .	20
2.1.1	Introduction . . . . .	20
2.1.2	1-D Discrete Dyadic Wavelet Transform Revisited . . . . .	20
2.1.3	Steerable Functions . . . . .	29
2.1.4	Steerable Dyadic Wavelet Transform . . . . .	31
2.1.5	Multiscale Spline Derivative-Based Transform . . . . .	33
2.1.6	Image Fusion Application . . . . .	37
2.1.7	Summary . . . . .	40
2.2	A Parallel Algorithm to Support Interactive Wavelet Processing on a Radiol- ogist Workstation . . . . .	42
2.2.1	Introduction . . . . .	42
2.2.2	Architectural Overview of SPARCstation SX System . . . . .	43
2.2.3	XIL Imaging Library . . . . .	44
2.2.4	An Algorithm for High-Speed Wavelet Analysis . . . . .	45
2.2.5	Summary . . . . .	54
2.3	A Continuous Scale Discrete Wavelet Transform for the Detection of Spicular Masses in Mammograms . . . . .	54
2.3.1	Introduction . . . . .	54
2.3.2	Definitions and Notation . . . . .	56
2.3.3	A Continuous Scale Discrete Wavelet Transform . . . . .	57
2.3.4	An Efficient Algorithm for Analysis and Synthesis . . . . .	61
2.3.5	Examples of CSDWT Filters . . . . .	62
2.3.6	Application of the CSDWT Algorithm . . . . .	64
2.3.7	Frequency Approximation vs Projection Methods . . . . .	73
2.3.8	Summary . . . . .	74

2.4	Enhancement of Mammograms from Oriented Information . . . . .	75
2.4.1	Introduction . . . . .	75
2.4.2	Background . . . . .	76
2.4.3	Methodology . . . . .	78
2.4.4	Experimental Results . . . . .	82
2.4.5	Summary . . . . .	84
2.5	Multivoice Undecimated Wavelet Transforms . . . . .	88
2.5.1	Introduction . . . . .	88
2.5.2	The Short-Time Fourier Transform . . . . .	88
2.5.3	The Wavelet Transform . . . . .	91
2.5.4	The À Trous and Mallat's Algorithms . . . . .	102
2.5.5	The À Trous Algorithm as an Exact Wavelet Transform . . . . .	110
2.5.6	The Sine-Gabor Wavelet . . . . .	113
2.5.7	The Sine-Gabor Wavelet Frame . . . . .	114
2.5.8	Summary . . . . .	119
2.6	Circular Mass Recognition Based on the Hough Transform . . . . .	120
2.6.1	Introduction . . . . .	120
2.6.2	Algorithm . . . . .	122
2.6.3	Results . . . . .	126
2.6.4	Synthetic Images . . . . .	128
2.6.5	Summary . . . . .	135
<b>3</b>	<b>Conclusions</b>	<b>137</b>
<b>A</b>	<b>Appendix</b>	<b>149</b>
A.1	Initial Condition . . . . .	149
A.2	Proof of Theorem 3 in Section 2.3.7 . . . . .	150
A.3	Derivation of Equation 179 . . . . .	151



# List of Figures

1	Configuration of output. . . . .	18
2	Filter bank implementation of a 1-D discrete dyadic wavelet transform. . . .	24
3	Comparison of two discrete implementations using $\text{sinc}(x)$ as an input. . . .	25
4	Wavelets $\psi(x) = \frac{d^d \beta_{p+d}(x)}{dx^d}$ for $p \in \{0, 1, 2\}$ and $d \in \{1, 2, 3\}$ . . . . .	28
5	Filter bank implementation of a spline derivative-based transform. . . . .	38
6	Image fusion examples. . . . .	41
7	SPARCstation SX system high-level architecture [39]. . . . .	43
8	Solaris foundation graphics libraries and layered interfaces [40]. . . . .	44
9	(a) Traditional linear convolution (denoted by $\otimes$ ) at scale $l = 1, l = 2$ . One zero is padded between non-zero filter coefficients in the filter kernel at scale 2. (b) Alternative approach that avoids zero-paddings between the non-zero filter coefficients. We first convolve the input image with the original filter kernel at the level 1. The convolved image is then divided into four sub-regions, marked in the background as <i>I</i> , <i>II</i> , <i>III</i> , and <i>IV</i> with half-grayed tone, each of which holds pixels that are precisely affected by the filter kernel coefficients at the next level $l = 2$ . Four distinct pixel shapes inside the input image are shown: $\bullet$ , $\triangle$ , $\square$ and $\times$ . Only pixels that have the same shape are processed by convolution at the next level (denoted by $\tilde{\otimes}$ ). . . . .	46
10	Multiresolution block convolution for image synthesis. Four sub-regions in the input image at a coarser level, marked as <i>I</i> , <i>II</i> , <i>III</i> , and <i>IV</i> , are merged into a larger region at a finer level. Then, convolution is applied to the merged region. . . . .	47
11	(a) Selected ROI containing an ill-defined mass. (b) DC components at levels $l = 1$ , $l = 2$ and $l = 3$ after applying <b>Algorithm 1</b> . (c) Processed ROI with suspicious area enhanced. . . . .	50
12	A user interface, <b>XEnh</b> , that implements our high-speed parallel algorithm for interactive enhancement. . . . .	51
13	RSNA '96 Exhibition Home-page ( <a href="http://www.iprg.cise.ufl.edu/demo">http://www.iprg.cise.ufl.edu/demo</a> ). . . . .	52
14	Hands-on section of the RSNA '96 Exhibition Home-page shown in Figure 13. (a) M019 case with biopsy proven masses. A ROI is preselected. (b) Result of enhancement on the ROI selected in (a). . . . .	53
15	Data acquisition processing. . . . .	57
16	Initial conditions. . . . .	58
17	initcond . . . . .	58
18	Filter bank for CSDWT decomposition and reconstruction for 3 levels of analysis. . . . .	62
19	Sample of fast decomposition algorithm of white noise at scale 14.25. . . . .	64

20	(a) X-ray image of the RMI156 phantom; (b) Coefficient minima obtained at scale 53.8; (c) Coefficient minima obtained at scale 32; (d) Coefficient minima obtained at scale 64; (e) Schematic representation of mammographic features within the phantom. . . . .	66
21	(a) Mammogram rml041; (b) Mammogram lcc046; (c) ROI cropped from lcc041; (d) ROI cropped from lcc046. . . . .	67
22	(a) Labeling of wavelet coefficient minima within selected ROI from rml041 (soft-thresholding applied). “ <i>JUNCTIONs</i> ” are shown in BLUE, “ <i>TERMINALs</i> ” in GREEN, and “ <i>CONTOURs</i> ” in red; (b) Labeling of wavelet coefficient minima within ROI from lcc046; (c) 4X magnification of the center cluster representing a mass shown in (a); (d) 4X magnification of the center cluster representing a mass shown in (b). . . . .	68
23	(a) Wavelet coefficients of rml041 at scale 124.8; (b) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass detected in rml041. Note consistency of measure for contours identified within each cluster; (c) Wavelet coefficients of lcc046 at scale 112; (d) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass shown in lcc046. . . . .	69
24	(a) Mammogram lcc002; (b) Mammogram lml015; (c) ROI cropped from lcc002; (d) ROI cropped from lml015. . . . .	70
25	(a) Labeling of wavelet coefficient minima within selected ROI from lcc002 (soft-thresholding applied). “ <i>JUNCTIONs</i> ” are shown in BLUE, “ <i>TERMINALs</i> ” in GREEN, and “ <i>CONTOURs</i> ” in red; (b) Labeling of wavelet coefficient minima within ROI from lml015; (c) 4X magnification of the center cluster representing a subtle mass shown in (a); (d) 4X magnification of the center cluster representing a subtle mass shown in (b). . . . .	71
26	(a) Wavelet coefficients of lcc002 at scale 126.4; (b) Color coding of a local texture measure computed normal to the contour of a spicular distortion around a mass shown in lcc002. Note consistency of measure for contours identified within each cluster; (c) Wavelet coefficients of lml015 at scale 89.6; (d) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass shown in lml015. . . . .	72
27	(a) Wavelet coefficients for lml015 at scale 64; (b) Wavelet minima in lml015.	73
28	Two-dimensional impulse responses of three basis functions. . . . .	80
29	Overview of processing for Steps 1–3. . . . .	81
30	A schematic diagram of nonlinear operator at level $i$ . . . . .	83

31	The ability of detecting objects oriented along distinct directions at various scales: (a) Original test image (mathematical phantom), (b) enhanced image after processing. . . . .	84
32	Mammogram with calcifications: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image. . . . .	85
33	Mammogram with a stellate lesion: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image. . . . .	86
34	Mammogram with a mass: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image. . . . .	87
35	Square of the product of the time and frequency resolution for the sine-Gabor wavelet as a function of $\sigma_0$ and $\omega_0$ . . . . .	116
36	Sine-Gabor wavelet (left) and magnitude of its Fourier transform (right) with (a) $\sigma_0 = 5/2$ , $\omega_0 = 1$ , (b) $\sigma_0 = 1$ , $\omega_0 = 1$ and (c) $\sigma_0 = 1.0657$ , $\omega_0 = 0.0299$ . . .	117
37	First derivative of a Gaussian (left) and the magnitude of its Fourier transform (right) with $\sigma_0 = 1.0657$ . . . . .	119
38	Demonstration of the fact that the center of circle $c$ , $O$ , belongs to the line that perpendicularly bisects the segment defined by points $A$ and $B$ (both points belong to the circumference of the circle). . . . .	121
39	Graphic illustrating our implementation of a 2D HT for circle center localization. Points $P_0$ and $P_3$ of the connected component $P_0, P_1, \dots, P_9$ , are assumed to belong to the same circle and all points that belong to the line that perpendicularly bisects the line segment they define receive votes. The same process is repeated for the pairs $P_1-P_4$ , $P_2-P_5$ etc. For this particular case the length of the window is equal to three. If one chooses to connect points $P_0-P_5$ , $P_1-P_6$ , etc., the length of the window is equal to five. . . . .	123
40	The radius histogram for the center of a circle. All bins get a small number of votes except for the bin to which the circle belongs. A sharp maximum appears for the bin that contains the circle. . . . .	124
41	All pixels belonging to a digital circle can be bounded by two concentric circles, $c_1$ and $c_2$ , whose radii differ by $h$ . For the digital circle of this figure, all pixels are spread between two neighboring bins, $b_1$ and $b_2$ . . . . .	125
42	If an inaccurate estimate of the center of the circle is provided by the first step, the radius histogram will give two peaks. In this example, the estimated center is point $P$ , while the true center is point $O$ . Obviously, these peaks will appear at bins with distance $r_1$ and $r_2$ from the estimated center. . . .	126

43	Comparison of the robustness $P$ as a function of the $SNR$ for three circles of radii: (a) $r = 20$ ; (b) $r = 30$ ; and (c) $r = 50$ pixels. As the step becomes larger $P$ increases. . . . .	127
44	Figure illustrating the key stages of the algorithm for a high $SNR$ synthetic image ( $SNR = 26$ ). (a) Original image with a circle at its center. The radius of the circle was 20 pixels; (b) edge detected image; (c) plotting of the voting space in a mesh format. The peak indicates the possible existence of a circle. The coordinates of the peak provide an estimation of the coordinates of the center of the circle; (d) histogram of the feature pixels as a function of their distance from the identified center (peak in the voting space); (e) filtered histogram. The identified peak was at $r = 20$ ; (f) detected circle superimposed on the original image. . . . .	129
45	Figure illustrating the key stages of the algorithm for a low $SNR$ synthetic image ( $SNR = 2$ ). (a) Original image with a circle at its center. The radius of the circle was 20 pixels; (b) edge detected image; (c) plotting of the voting space in a mesh format. The peak indicates the possible existence of a circle. The coordinates of the peak provide an estimation of the coordinates of the center of the circle; (d) histogram of the feature pixels as a function of their distance from the identified center (peak in the voting space); (e) filtered histogram. The identified peak was at $r = 20$ ; (f) detected circle superimposed on the original image. . . . .	131
46	Plots of the mean value of the maximum value of the filtered histogram $h_{max}$ , for 20 experiments, as a function of the $SNR$ . Continuous lines denote results using the bisection-based HT, while dashed lines denote results of the gradient-based HT, for circles of radii: (a) $r = 20$ ; (b) $r = 30$ ; and (c) $r = 50$ . While for high $SNR$ (i.e. $SNR > 20$ ) the gradient-based HT works better, for low $SNR$ drop the bisection-based HT gives better results. . . . .	132
47	Plots of the error in the determination of the center of the circle with the use of the bisection-based method (continuous lines) and the gradient-based method (dashed lines) for circles of radii: (a) $r = 20$ ; (b) $r = 30$ ; and (c) $r = 50$ . For most of the cases the error of the gradient-based method is larger than the error of the bisection-based method. As shown in these plots, the error of the bisection-based method for a wide range of noise levels remained smaller than a pixel. . . . .	133

48	Figure illustrating the steps of the algorithm on a digitized phantom. (a) Image of a phantom; (b) output of Canny edge detection; (c) detected circles superimposed on the original image. . . . .	134
49	Figure illustrating the steps of the algorithm on a digitized mammogram. (a) Digitized mammogram; (b) output of Canny edge detection; (c) detected circles superimposed on the original image. The algorithm correctly detected the two overlapping masses while giving a false positive (top middle of the image). . . . .	136
50	Modify initial condition . . . . .	149

# 1 Introduction

Breast cancer is the most frequently diagnosed malignancy among women in the United States [1]. In 1996 the American Cancer Society estimates that 184,300 women will be newly diagnosed with breast cancer and that 44,300 will die from the disease [1]. Breast cancer accounts for 31% of all cancers detected and 17% of all cancer deaths, and ranks as the second leading cause of death from cancer among women in the United States [1]. Five year survival rates are generally very high (93%) for breast cancer staged as being localized, falling to 72% for regional disease and only 18% for distant disease [2]. The early detection of breast cancer is clearly a key ingredient of any strategy designed to reduce breast cancer mortality.

Mammography's role is the early detection of breast cancer. Although more accurate than any other modality, existing techniques for mammography only find 80 to 90% of the breast cancers. Moreover, in 7 to 10% of cases, the cancer will not be visible on the mammogram. It has been suggested that mammograms as normally viewed, display only about 3% of the total information detected. Perception is a problem particularly for patients with dense fibroglandular patterns. The importance of diagnosis of breast cancer at an early stage is critical to patient survival. The general inability to detect small tumors and other salient features within mammograms motivates our investigation.

The goal of this project is to develop an interactive diagnostic tool for radiologists that shall refine the perception of mammographic features (including lesions, masses and calcifications) and improve the accuracy of diagnosis. By improving the visualization of breast pathology we can increase the chances of early detection of breast cancers (improve quality) while requiring less time to evaluate mammograms for most patients (lower costs). We are investigating a methodology for accomplishing mammographic feature analysis through multiresolution representations. We have shown that efficient (nonredundant) representations may be identified from digital mammography and used to reconstruct and enhance specific mammographic features within a continuum of scale space. Such "focused" reconstructions may complement existing modalities and allow a radiologist to examine interactively diagnostic features within a selected scale space. Similar to traditional coarse to fine matching strategies, the radiologist may first choose to look for coarse features (e.g. dominant mass) within low frequency levels of a wavelet transform and later examine finer features (e.g. microcalcifications) at higher frequency levels. In addition, features may be extracted by applying geometric constraints within each level of the transform. Choosing wavelets (or analyzing functions) that are simultaneously localized in both space and frequency, results in a powerful methodology for image analysis.

A major reason for poor visualization of small malignant masses is the subtle difference in x-ray attenuation between normal glandular tissues and malignant disease [3]. This fact makes the detection of small malignancies problematical, especially in younger women who have denser breast tissue. Although calcifications have high inherent attenuation properties, their small size also results in a low subject contrast [4]. As a result, the visibility of small tumors, and any associated microcalcifications, will always be a problem in mammography as it is currently performed using analog film.

In this report, we describe some exciting results accomplished during the past year. Below we describe in executive summary, progress related to the specific tasks and objectives stated in Phases III and IV of our original Statement of Work.

In the sections below, we describe in overview, our processing algorithms, experimental methods and sample results obtained. In addition, we list in summary, publications of our researchers accomplished during the past year of our investigation. Finally, we include our response to the reviewers comments and suggestions regarding contractual and technical issues in reference to last year's report.

## **1.1 Overview of Contents**

### **Generalizing the Discrete Dyadic Wavelet Transform**

Traditional orthogonal and biorthogonal wavelet representations may introduce aliasing and anisotropies due to the lack shift and rotation invariance. In screening mammography intolerance to translation and rotation is clearly not desirable. The multiscale spline derivative-based transform presented in this section is a redundant representation which does not introduce such artifacts.

As preliminary results, two examples of image fusion are presented. Original mammograms were fused with their globally enhanced counterparts to obtain enhanced images having an appearance more similar to the original mammograms. The fused images appears easier for radiologists to interpret without additional training on wavelet processing techniques.

### **A Parallel Algorithm to Support Interactive Wavelet Processing on a Radiologist Workstation**

We have designed an efficient parallel algorithm to support interactive analysis and multiscale processing. We show how this *multiscale block* algorithm works and demonstrate the amount of speedup achieved over traditional linear convolution schemes and conventional FFT methods.

## **A Continuous Scale Discrete Wavelet Transform for the Detection of Spicular Masses in Mammograms**

We introduce a simple method to evaluate continuous wavelet transform (CWT). We present a fast decomposition and reconstruction algorithm to perform 2-D analysis at arbitrary scales. Arbitrary scale analysis is shown to be critical for optimal feature detection and enhancement in mammograms [5]. Our algorithm was able to detect a mass that could not be seen using conventional windowing and leveling or traditional contrast enhancement methods.

### **Enhancement of Mammograms from Oriented Information**

Mammograms depict all the significant changes in breast disease. The primary radiographic signs of cancer are related to tumor mass, its density, size, shape, borders and calcification content. These features may be extracted according to their coherence and orientation. Such features provide important visual cues for radiologists to locate suspicious areas easily. In this investigation, an artifact free enhancement algorithm based on multiscale wavelet analysis is presented. First an image was decomposed using a fast wavelet transform algorithm. At the same time, the energy and phase information at each level were determined using a set of separable steerable filters. Then a measure of coherence within each level was obtained by weighting the energy with the ratio of projections of the energy within a specified window onto the central point of the window with respect to the total energy within each window. Finally, a nonlinear operation, integrating coherence and orientation information, was applied to modify the transform coefficients. These modified coefficients were then reconstructed, via an inverse fast wavelet transform, resulting in an improved visualization of mammographic features. The novelty of this algorithm lies in its detection of directional features and removal of unwanted perturbations. Compared to existing multiscale enhancement approaches, these results appear more familiar to radiologists and naturally close to the original mammogram.

### **Multivoice Undecimated Wavelet Transforms**

For many analysis/synthesis applications the sampling provided by the traditional wavelet transform is inadequate. In particular, a finer grid is needed when computing the wavelet transform for applications that require "good" time-frequency localization. This can be achieved by computing an undecimated wavelet transform and the addition of voices. In this section, we review the tools that will allow us to compute an exact multivoice undecimated wavelet transform for an arbitrary wavelet. Furthermore, we introduce a wavelet function that exhibits nearly optimum time-frequency localization (the sine-Gabor wavelet) and show that it is indeed a wavelet frame. However, having a sine-Gabor wavelet



with nearly optimum time-frequency localization and obtaining a tight frame at the same time is not possible. Both constraints are important because they allow for “good” time-frequency localization and efficient reconstruction, respectively. Obtaining a better (tighter) frame can be achieved by adding voices. This not only eases the reconstruction but is also a requirement for applications in time-frequency analysis. Our current work is concerned with extending the sine-Gabor wavelet to multivoice framework and with the computation of multivoice undecimated wavelet transforms. Our future work includes wavelet reconstruction in the multivoice/undecimated framework as well as extending our results to the analysis of two dimensional data for mammography.

### **Circular Mass Recognition Based on the Hough Transform**

In this chapter we present a two-step algorithm for the recognition of circular masses. The first step uses a 2D Hough Transform for the detection of the centers of possible circular shapes and the second step validates their existence by radius histogramming. The 2D Hough Transform described in this section makes use of the property that every chord of a circle passes through its center. We present sample results from experiments using both synthetic and real data demonstrating that our method is more robust to noise and complex backgrounds (typically found in real mammograms) than standard gradient based methods. The promise of the method is demonstrated with its application on a digitized phantom containing four circular masses and on a digital mammograms.

## **1.2 Publications**

Below, we list in summary, an update of publications accomplished during 1996.

1. D. Chen, A. Laine, J. G. Harris, and W. Huda, “A continuous scale discrete wavelet transform for the detection of spicular masses in mammograms,” submitted to *IEEE Transactions on Signal Processing*, February 1997.
2. D. Ioannou, W. Huda, A. F. Laine, and I. Koren, “Enhancement of computer simulated masses for mammography via wavelet analysis,” submitted to *IEEE Transactions on Medical Imaging*, January 1997.
3. I. Koren and A. Laine, “A discrete dyadic wavelet transform for multidimensional feature analysis,” in M. Akay (Editor), *Time-Frequency and Wavelet Transforms in Biomedical Engineering*, New York, NY: IEEE Press, 1997.
4. S. Schuler and A. Laine, “Hexagonal QMF banks and wavelets,” in M. Akay (Editor), *Time-Frequency and Wavelet Transforms in Biomedical Engineering*, New York, NY: IEEE Press, 1997.

5. J. Fan and A. Laine, "Multiscale contrast enhancement and denoising in digital radiographs," in A. Aldroubi and M. Unser (Editors), *Wavelets in Medicine and Biology*, Boca Raton, FL: CRC Press, 1996, pp. 163–189.
6. A. F. Laine and X. Zong, "A multiscale sub-octave wavelet transform for de-noising and enhancement," in *Wavelet Applications in Signal and Image Processing IV, Proceedings of SPIE*, Denver, CO, August 6-9, 1996, vol. 2825. pp. 238–249.
7. A. Laine, W. Huda, D. Chen, and J. Harris, "Segmentation of masses using continuous scale representations," in *Digital Mammography '96, Proceedings of the 3rd International Workshop on Digital Mammography*, Chicago, U.S.A., 9-12 June 1996, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Editors, Amsterdam, The Netherlands: Elsevier, 1996, pp. 447–450.
8. I. Koren, A. Laine, F. Taylor, and M. Lewis, "Interactive wavelet processing and techniques applied to digital mammography," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Atlanta, GA, May 7-10, 1996, vol. 3, pp. 1415–1418.
9. X. Zong, A. F. Laine, E. A. Geiser, and D. C. Wilson, "De-noising and contrast enhancement via wavelet shrinkage and nonlinear adaptive gain," in *Wavelet Applications III, Proceedings of SPIE*, Orlando, FL, April 8-12, 1996, vol. 2762, pp. 566–574.
10. D. Chen, J. G. Harris, and A. Laine, "Automatic scale detection," in *Visual Communications and Image Processing, Proceedings of SPIE*, Orlando, FL, March 17-20, 1996, vol. 2727, pp. 960–972.
11. G. Qu, W. Huda, and C. J. Belden, "Comparison of trained and untrained observers using subjective and objective measures of imaging performance," in *Academic Radiology*, 1996, vol. 3, pp. 31–35.

### 1.3 Responses to Technical and Contractual Issues

We believe this report shows significant progress with respect to the objectives described in Phase's III and IV of our original Statement of Work (SOW). We acknowledge the reviews concerns regarding the performance of our methods in the context of complex backgrounds (typical of real mammograms) and generating false positives. Sections 1.1, 1.3, 1.4, 1.6 describe processing techniques that are artifact free and significantly minimize the generation of false positives compared to previous methods. We provide experimental evidence and theoretical support to justify these claims in each case.

In addition, with regard to the ROC studies mentioned in Phase V of our SOW, we include

below a description of our ongoing investigation. It is expected that these studies shall be complete within three months time.

## **1.4 Study in Progress: An ROC Comparison Between Digital Mammography and Screen-Film Using an Anthropomorphic Breast Phantom**

### **1.4.1 Introduction**

Several studies have been performed comparing the performance of digital mammography with screen-film mammography with ambiguous results. In one case, the conspicuity of low contrast lesions on a flat grey background was greater using digital acquisition and display than using screen-film combinations. (1) In another case, a similar research protocol was used with simulated masses placed on an anthropomorphic breast phantom, resulting in greater conspicuity using the screen-film combination than with the digital methods. (2) In spite of the conflicting results, the studies provided useful for the following reasons. First, we were able to perfect the methodology for merging simulated masses with the realistic background of the anthropomorphic breast phantom, and second, the ROC methodology developed indicated that we could get valid results using individuals with a wide range of experience. The abilities to generate images with the **truth** established and to evaluate the images without the requirement for sub-specialty radiologist readers enable us to use a faster, less costly approach to applying established ROC methodology to the evaluation of our processed images.

The protocol developed by Qu, Huda, et.al. [6] will be modified slightly to be used to evaluate steerable wavelet processing of digital mammograms as described in Chapter 2.4. In order to evaluate and exploit the power of the steerable wavelets, two shapes of simulated masses will be used. The first will simulate a smooth, mostly round mass, while the second will simulate an oblong mass with rather irregular edges. The purpose of the study is to evaluate the effectiveness of using the steerable wavelets on different shaped simulated masses in the context of complex backgrounds. Since we will be using digital enhancement methods, only digital mammography images will be used for this study.

### **1.4.2 Method**

#### **Phantoms**

Digital images will be generated of the central region of an RMI 165 anthropomorphic phantom. A 1 mm thick acrylic base will be positioned on top of the RMI 165 phantom containing a  $3 \times 3$  array of 1.67 cm  $\times$  1.67 cm regions of interest (ROI). Radiopaque

markers will be located in the four corners of each ROI to permit these regions to be identified in the resultant radiographic images. Each image will be 5 cm  $\times$  5 cm in size and will contain 9 identifiable ROIs.

The test objects are designed to simulate masses in a mammogram. These simulated masses will be constructed of a 300 microns (t) film polymer and positioned in the central region of each ROI with a randomly selected orientation. Two types of simulated masses will be made, the first will simulate a smooth, mostly round, mass with a diameter of approximately 1.5 cm, while the second will simulate an oblong mass with rather irregular edges with a length of approximately 1.5 cm and a width of approximately 0.5 cm. Each of the masses will be made with three thicknesses, 5t, 4t, and 3t (six total mass test objects). Nine radiographs with different inserts were made with the simulated masses located in randomly selected ROIs for a total of 81 ROIs. Approximately half of the ROIs will contain a test object and the six test objects will be evenly distributed randomly throughout the images. In addition, the test objects will be imaged on a uniform background to be used to demonstrate the expected shapes to the readers.

### **Mammographic Imaging**

Digital mammographic images of the RMI 165 (+ insert) will be generated using the LoRad DSM using 28 kVp/72 mAs as technique factors. The images will be transferred by disk to a sun workstation where wavelet processing will be applied. The nine images (each with nine ROIs) will then be merged to form a single image with a 156 pixel black border and 100 pixel black separators between the nine images resulting in a 2048 pixel  $\times$  2048 pixel image (see Figure 1). Using tools developed at the University of Florida, the image will be converted to a DICOM image and will be transferred to a high resolution workstation. There will be two images generated with phantom data to contain unprocessed image data and wavelet processed image data. Each 512 pixel block will either be unprocessed or processed and the blocks will be randomly placed on the image. The single flat background image with the six test objects will also be converted to a 2048 pixel  $\times$  2048 pixel image by surrounding the image with a 768 pixel black border.

### **Workstation Display**

The high resolution DICOM workstation developed at the University of Florida is used for diagnostic interpretation of scanned radiographs and computed radiography. The megascan monitors and frame buffers are driven by a Sun Sparc20 computer with image display software written in-house. Images are displayed at full 2048  $\times$  2048 resolution and can be

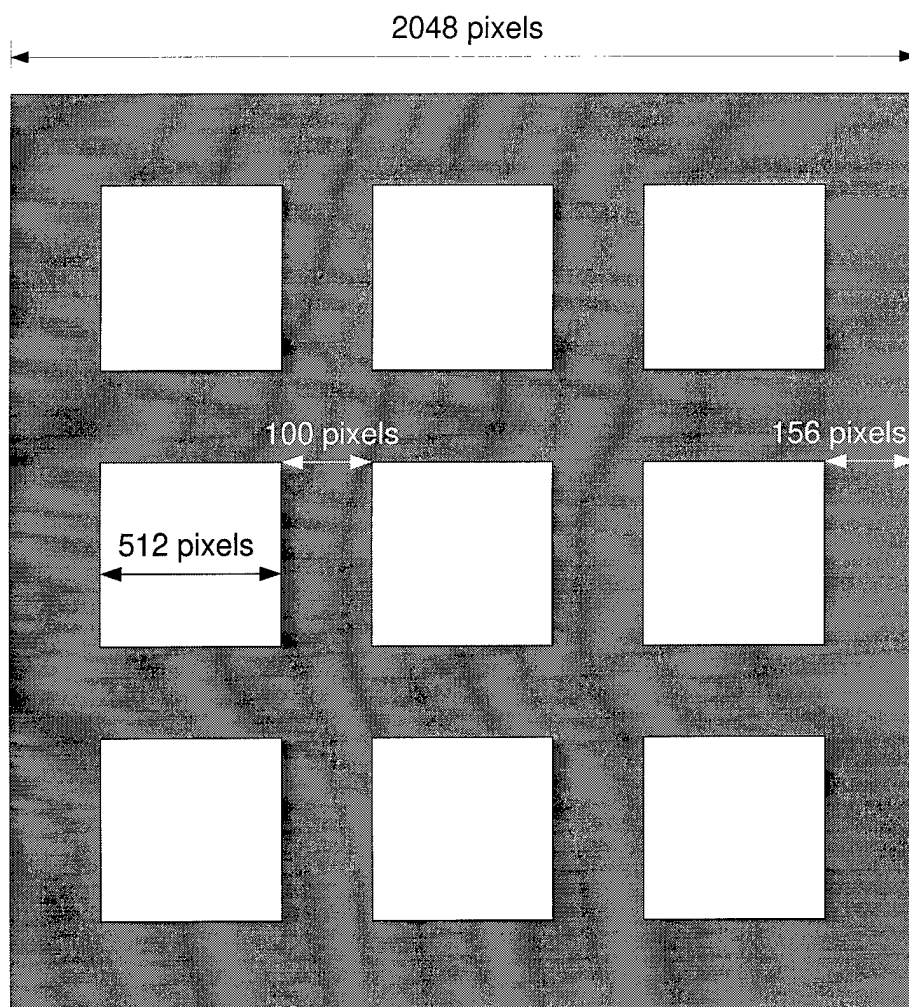


Figure 1: Configuration of output.

manipulated with zoom / roam, window / level, and by inverting the grey scale.

### Reader Performance

Readers will be shown the phantom images on one of the two high resolution monitors on the megascan with the flat grey image demonstrating the test objects on the other. By making all three images 2048 pixels  $\times$  2048 pixels, we are assured that the images will not be subsampled or enlarged, since this is the optimal image size for the workstation. We will use 12 readers, six radiologists (1 mammographer and 5 senior radiology residents) and six scientists experienced in reviewing medical physics images (two basic scientists and four graduate students) to score each ROI. The readers will be asked to review each ROI and indicate the likelihood of the ROI containing a simulated mass using a scale ranging from 0

(test object definitely absent) to 10 (test object definitely present). A rank of 5 indicates that there is a 50% probability of the simulated mass being present.

The results for each reader and processed / unprocessed ROI will be fitted to a standard ROC curve, which plots the True-Positive Fraction (TPF) vs False-Positive Fraction (FPF) as the detection threshold is varied. The ROCFIT software package (Metz, University of Chicago) will be used to fit the acquired True-Positive Fraction (TPF) vs False-Positive Fraction (FPF) data and to obtain the resultant area,  $A_z$ , under each ROC curve. A paired t-test analysis will be performed to assess the significance of the differences in the area under the ROC curve.

## 2 Body

### 2.1 Generalizing the Discrete Dyadic Wavelet Transform

#### 2.1.1 Introduction

Traditional orthogonal and biorthogonal wavelet representations, now ubiquitous in image processing and computational vision, lack shift and rotation invariance, and may introduce aliasing and anisotropies. Such artifacts are especially problematic for medical image analysis, but can be eliminated by the use of redundant representations.

Since a mammogram can easily be translated or rotated, it is not desirable that the result of processing be significantly affected by positioning during acquisition or digitization. The two-dimensional transforms presented in this section do not introduce artifacts due to translation and rotation invariance.

First, the one-dimensional discrete dyadic wavelet transform [7] was generalized to higher-order derivatives, equipped with a better initialization procedure, and used for derivation of a steerable dyadic wavelet transform. The steerable dyadic wavelet transform is both translation and rotation invariant, but exact reconstruction is problematic via implementation in the spatial domain. This was solved by a spline-based approximation which was used for a multiscale spline derivative-based transform.

A multiscale spline derivative-based transform was implemented in the spatial domain as a filter bank consisting of  $x$ - $y$  separable filters only. In addition to not introducing artifacts, the transform enables directional analysis of images across dyadic scales.

As a sample application, two examples of image fusion are presented. Original mammograms were fused with an enhanced version to obtain an enhancement that looked more familiar to radiologists than mammograms, enhanced by global methods previously developed.

#### 2.1.2 1-D Discrete Dyadic Wavelet Transform Revisited

In our previous report, we presented a complete description of discrete dyadic wavelet transform. For clarity, we summarize below definitions of the continuous and discrete transform [7].

The dyadic wavelet transform of a function  $s(x) \in L^2(\mathbf{R})$  is defined as a sequence of functions

$$\{W_m s(x)\}_{m \in \mathbf{Z}}, \quad (1)$$

where  $W_m s(x) = s * \psi_m(x) = \int_{-\infty}^{\infty} s(t) \psi_m(x-t) dt$ , and  $\psi_m(x) = 2^{-m} \psi(2^{-m}x)$  is a wavelet  $\psi(x)$  expanded by a dilation parameter (or scale)  $2^m$ . To ensure coverage of the frequency

axis the requirement on the Fourier transform of  $\psi_m(x)$  is the existence of  $A_1 > 0$  and  $B_1 < \infty$  such that

$$A_1 \leq \sum_{m=-\infty}^{\infty} |\hat{\psi}(2^m \omega)|^2 \leq B_1$$

is satisfied almost everywhere. The constraint on the Fourier transform of the (nonunique) reconstructing function  $\chi(x)$  is

$$\sum_{m=-\infty}^{\infty} \hat{\psi}(2^m \omega) \hat{\chi}(2^m \omega) = 1.$$

A function  $s(x)$  can then be completely reconstructed from its dyadic wavelet transform using the identity

$$s(x) = \sum_{m=-\infty}^{\infty} W_m s * \chi_m(x),$$

where  $\chi_m(x) = 2^{-m} \chi(2^{-m} x)$ .

In numerical applications, processing is performed on discrete rather than continuous functions. When the function to be transformed is in the discrete form, the scale  $2^m$  can no longer vary over all  $m \in \mathbf{Z}$ . Finite sampling rate prohibits the scale from being arbitrarily small, while computational resources restrict the use of an arbitrarily large scale. Let the finest scale be normalized to 1 and the coarsest scale set to  $2^M$ .

The smoothing of a function  $s(x) \in L^2(\mathbf{R})$  is defined as

$$S_m s(x) = s * \phi_m(x),$$

where  $\phi_m(x) = 2^{-m} \phi(2^{-m} x)$  with  $m \in \mathbf{Z}$ , and  $\phi(x)$  is a smoothing function (i.e., its integral is equal to 1 and  $\phi(x) \rightarrow 0$  as  $|x| \rightarrow \infty$ ).

In [7], a real smoothing function  $\phi(x)$  was selected, whose Fourier transform satisfied

$$|\hat{\phi}(\omega)|^2 = \sum_{m=1}^{\infty} \hat{\psi}(2^m \omega) \hat{\chi}(2^m \omega). \quad (2)$$

In addition, it was shown that any discrete function of finite energy  $s(n) \in l^2(\mathbf{Z})$  can be written as the uniform sampling of some function smoothed at scale 1, i.e.,  $s(n) = S_0 f(n)$ , where  $f(x) \in L^2(\mathbf{R})$  is not unique. Thus, the discrete dyadic wavelet transform of  $s(n)$  for any coarse scale  $2^M$  was defined as a sequence of discrete functions

$$\{S_M f(n + s), \{W_m f(n + s)\}_{m \in [1, M]}\}_{n \in \mathbf{Z}},$$



where  $s$  is a  $\psi(x)$  dependent sampling shift.

The above initialization  $s(n) = S_0 f(n)$  is rather standard in the discrete wavelet transform computation [8], although it yields correct results (i.e., the discrete wavelet transform is equal to the samples of its continuous counterpart) only when  $s(n) = S_0 s(n)$ . Here, we will concentrate on wavelets which are derivatives of spline functions and this will lead us to a simple initialization procedure [9] that alleviates the above problem.

For a certain choice of wavelets, the discrete dyadic wavelet transform can be implemented within a fast hierarchical digital filtering scheme. Next, we shall summarize the relations between filters, wavelets, and smoothing functions.

First, let us introduce a real smoothing function  $\varphi(x)$  such that (2) can be rewritten as<sup>1</sup>

$$\hat{\phi}(\omega) \hat{\varphi}(\omega) = \sum_{m=0}^{\infty} \hat{\psi}(2^m \omega) \hat{\chi}(2^m \omega), \quad (3)$$

and let us set

$$\phi(x) = \beta_p(x) = \frac{1}{p!} \sum_{i=0}^{p+1} (-1)^i \binom{p+1}{i} \left(x + \frac{p+1}{2} - i\right)^p u\left(x + \frac{p+1}{2} - i\right), \quad (4)$$

where  $\beta_p(x)$  denotes a central B-spline of order  $p$ . With the choice (4), we restrict ourselves to wavelets which are spline functions.

Computing (3) for the finest two scales shows that

$$\hat{\psi}(\omega) \hat{\chi}(\omega) = \hat{\beta}_p(\omega) \hat{\varphi}(\omega) - \hat{\beta}_p(2\omega) \hat{\varphi}(2\omega). \quad (5)$$

$\hat{\beta}_p(2\omega)$  can be related to  $\hat{\beta}_p(\omega)$  by expressing  $\hat{\beta}_p(2\omega)$  as (cf. Proposition 1 of [9])

$$\hat{\beta}_p(2\omega) = \frac{1}{2^{p+1}} \left( \frac{\sin(\omega)}{\sin(\frac{\omega}{2})} \right)^{p+1} \left( \frac{\sin(\frac{\omega}{2})}{\frac{\omega}{2}} \right)^{p+1},$$

and using  $\sum_{m=0}^M e^{j(m\omega+\theta)} = \frac{\sin(\frac{(M+1)\omega}{2})}{\sin(\frac{\omega}{2})} e^{j(\frac{M\omega}{2}+\theta)}$ :

$$\hat{\beta}_p(2\omega) = \left( \cos\left(\frac{\omega}{2}\right) \right)^{p+1} \hat{\beta}_p(\omega). \quad (6)$$

(Note that a relation similar to (6) can be derived for integer scales provided that the dilation parameter and the order  $p$  are not both even [9].)

Let  $F(\omega)$  be a digital filter frequency response and let  $F_s(\omega) = e^{j\omega s} F(\omega)$ .

---

<sup>1</sup>Note that the sum index determines the range of scales of the discrete transform: using (2) we have  $\hat{\psi}(2\omega)$  and  $\hat{\chi}(2\omega)$  at the finest scale of the transform, while for (3) we get  $\hat{\psi}(\omega)$  and  $\hat{\chi}(\omega)$ .

If we choose

$$\hat{\psi}(\omega) = G_{-s}(\omega) \hat{\beta}_p(\omega), \quad (7)$$

$$\hat{\varphi}(2\omega) = L_s(\omega) \hat{\varphi}(\omega), \quad (8)$$

$$\hat{\chi}(\omega) = K_s(\omega) \hat{\varphi}(\omega), \quad (9)$$

and

$$H(\omega) = e^{j\omega s} \left( \cos \left( \frac{\omega}{2} \right) \right)^{p+1}, \quad (10)$$

where  $s \in \{0, \frac{1}{2}\}$  is a filter dependent sampling shift needed for  $g(n)$ ,  $l(n)$ ,  $k(n)$ , and  $h(n)$  to be FIR filters, and insert Equations (6)–(10) into (5), we observe the relation between frequency responses of the filters

$$G(\omega)K(\omega) + H(\omega)L(\omega) = 1. \quad (11)$$

Similar to orthogonal and biorthogonal discrete wavelet transforms, the discrete dyadic wavelet transform can be implemented within a hierarchical filtering scheme. To derive such a digital filtering scheme, let us assume that  $\hat{s}(\omega)$  from (1) is bandlimited to  $[-\pi, \pi]$ . Using Shannon's sampling theorem [10] and (7) in the definition of the dyadic wavelet transform (1) with  $m=0$  shows

$$W_0 s(x) = \int_{-\infty}^{\infty} \sum_{i=-\infty}^{\infty} s(i) \text{sinc}(t-i) \sum_{m=-\infty}^{\infty} g_{-s}(m) \beta_p(x-t-m) dt.$$

By making use of the fact that the cardinal spline functions  $\eta_r(x)$  tend to the sinc function as their order  $r$  approaches infinity [11], and employing

$$\eta_r(x) = \sum_{i=-\infty}^{\infty} b_r^{-1}(i) \beta_r(x-i), \quad (12)$$

where  $b_r^{-1}(n)$  denotes a direct B-spline filter [12], we can write

$$\widehat{W_0 s}(\omega) \simeq S(\omega) B_r^{-1}(\omega) \hat{\beta}_r(\omega) \hat{\beta}_p(\omega) G_{-s}(\omega),$$

or, by using (6) and (10),

$$\mathcal{F}\{W_m s(x)|_{x=n}\} \simeq S(\omega) B_r^{-1}(\omega) B_{p+r+1}(\omega) G_{-s}(2^m \omega) \prod_{i=0}^{m-1} H_{-s}(2^i \omega). \quad (13)$$

Equation (13) entirely specifies the discrete dyadic wavelet transform decomposition, while the reconstruction follows from (5)–(10). Three levels of a filter bank implementation are

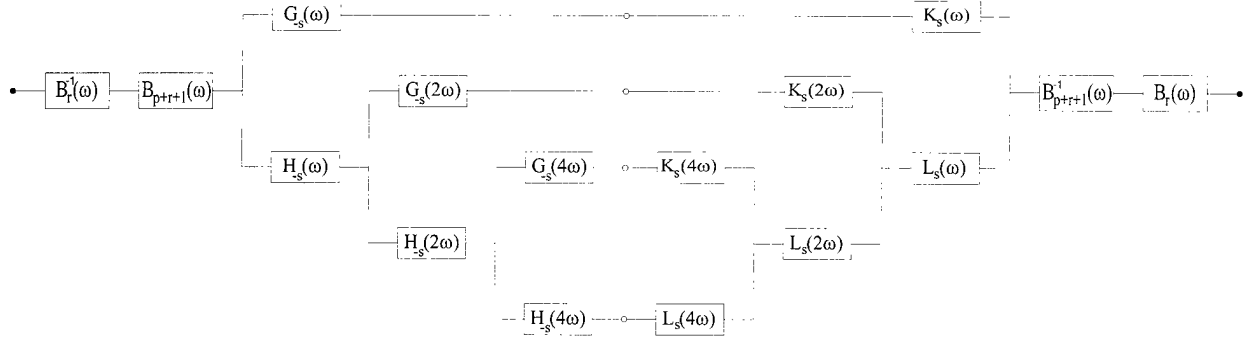


Figure 2: Filter bank implementation of a one-dimensional discrete dyadic wavelet transform decomposition (left) and reconstruction (right) for three levels of analysis.

shown in Figure 2. (Note that the initialization is the same as the one proposed in [9] and that except for the prefiltering and postfiltering, this scheme is implementing “algorithme à trous” [13].) Noninteger shifts at scale 1 for filters with  $s = \frac{1}{2}$  are rounded to the nearest integer.

We will now perform a simple experiment which will illustrate the difference between the implementation of the discrete dyadic wavelet transform as originally proposed in [7] (i.e., without prefiltering and postfiltering) and the one from Figure 2.

Let  $s(x) = \text{sinc}(x)$ ,  $p = 2$ , and  $g(n) = 2\delta(n+1) - 2\delta(n)$  (this particular choice for  $p$  and  $g(n)$  results in the same wavelet as was used by Mallat and collaborators [14, 7]). The dyadic wavelet transform of  $s(x)$  at a scale  $2^m$  (1) in the frequency domain is then

$$\widehat{W_m s}(\omega) = G_{-s}(2^m \omega) \hat{\beta}_2(2^m \omega) \text{rect}\left(\frac{\omega}{2\pi}\right). \quad (14)$$

The Fourier transform of the discrete dyadic wavelet transform of  $s(n) = \delta(n)$  at a scale  $2^m$  using spline based initialization follows from (13)

$$\mathcal{F}\{\widetilde{W_m s}(n)\} = B_r^{-1}(\omega) B_{r+3}(\omega) G_{-s}(2^m \omega) \prod_{i=0}^{m-1} H_{-s}(2^i \omega), \quad (15)$$

while the one using the algorithm from [7] equals

$$\mathcal{F}\{\widetilde{W_m s}(n)\} = G_{-s}(2^m \omega) \prod_{i=0}^{m-1} H_{-s}(2^i \omega). \quad (16)$$

In Figure 3 a comparison of magnitudes of (15) and (16) versus (14) is shown: in Figure 3(a) magnitudes of (14) (solid) and (16) (dashed) are plotted for  $m \in \{0, 1, 2, 3\}$ , while the dashed curves in 3(b) represent magnitudes of (15) with  $r=5$ .

By choosing the appropriate order  $r$ , (15) can approximate (14) in the interval  $[-\pi, \pi]$  arbitrarily good, while originally proposed (16) has troubles at finer scales. Mallat and

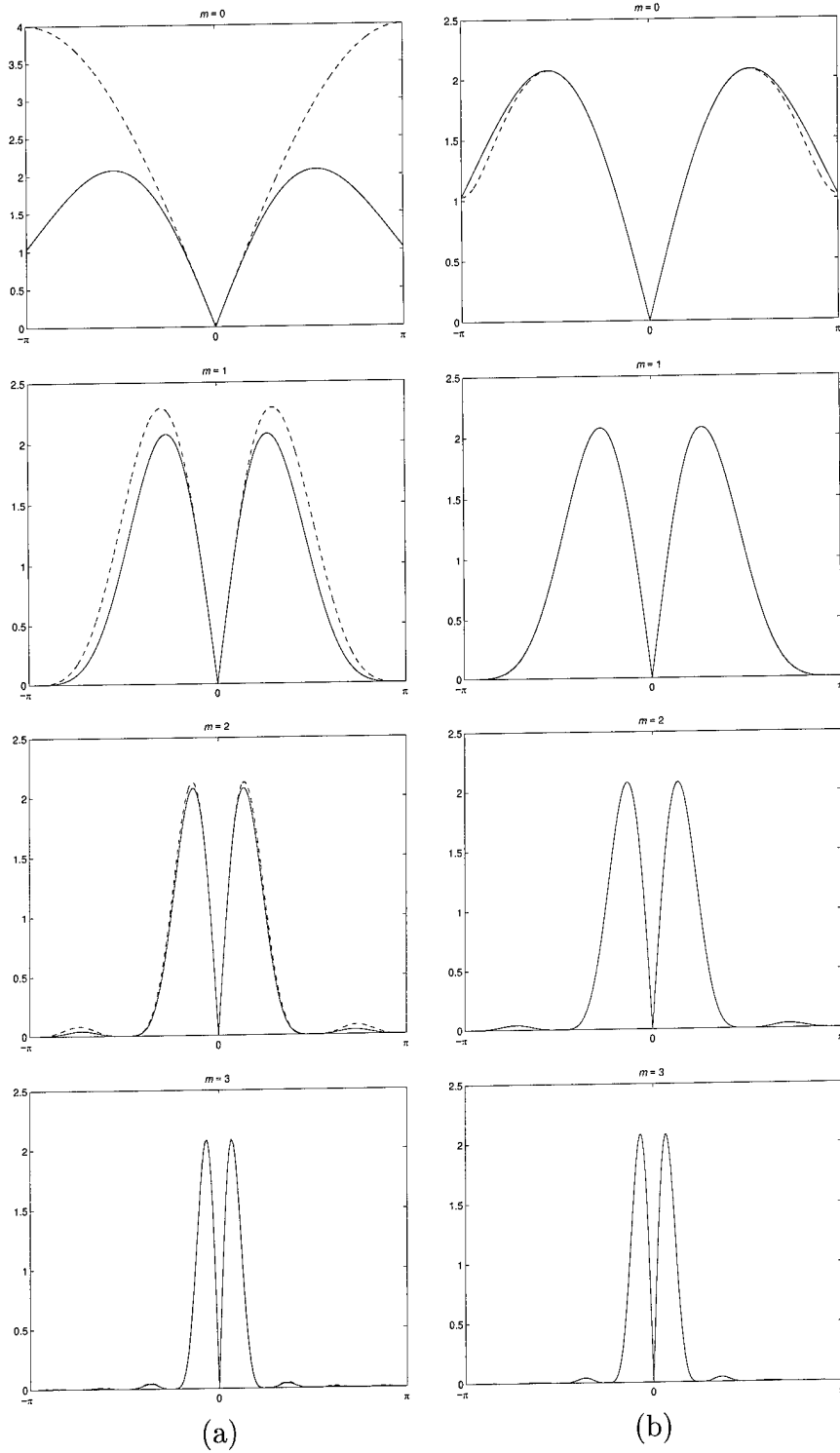


Figure 3: (a) Fourier transform magnitudes of the dyadic wavelet transform of  $s(x) = \text{sinc}(x)$  (solid) and the originally proposed discrete dyadic wavelet transform [7] of  $s(n) = \delta(n)$  (dashed). (b) Fourier transform magnitudes of the dyadic wavelet transform of  $s(x)$  (solid) and the discrete dyadic wavelet transform using quintic splines for interpolation of  $s(n)$  (dashed).

Zhong [7] noticed that there was a problem with their discrete transform computation, and introduced a set of constants associated with the discrete transform coefficients at dyadic scales. They chose the values of constants such that the transform coefficient modulus maxima remained constant over all dyadic scales for a step edge input signal. In relation to Figure 3(a) this is equivalent to multiplying  $\mathcal{F}\{\widetilde{W}_m s(n)\}$  by a distinct constant for each  $m$ . Clearly, this is an improvement over the situation depicted by Figure 3(a), but remains inferior to spline based initialization.

Next, we choose filters in the filter bank implementation of the discrete dyadic wavelet transform. As previously mentioned, we are interested in wavelets which are derivatives of spline functions. From the property of the central B-spline functions [15]

$$\frac{\partial \beta_p(x)}{\partial x} = \beta_{p-1}\left(x + \frac{1}{2}\right) - \beta_{p-1}\left(x - \frac{1}{2}\right),$$

it follows that  $G(\omega)$  in (7) is the Fourier transform of the difference operator centered around  $-s$ :

$$G(\omega) = e^{j\omega s} \left(2j \sin\left(\frac{\omega}{2}\right)\right)^d, \quad (17)$$

where  $d$  is the order of the derivative and the sampling shift for this filter is  $s = \frac{d \bmod 2}{2}$ .

Since  $H(\omega)$  was already given by (10), the remaining two filters to be determined are  $L(\omega)$  and  $K(\omega)$ . Both of them are (as is true for  $\varphi(x)$  and  $\chi(x)$ ) nonunique.

If we choose  $L(\omega)$  such that we can express  $K(\omega)$  in terms of a finite geometric series having the smallest number of elements for an arbitrary  $p$ , we get

$$L(\omega) = e^{-j\omega s} \sum_{m=1}^{\lfloor \frac{d+1}{2} \rfloor} (-1)^{m+1} \binom{\lfloor \frac{d+1}{2} \rfloor}{m} \left(\cos\left(\frac{\omega}{2}\right)\right)^{(p+1)(2m-1)} \quad (18)$$

and

$$K(\omega) = \frac{1}{(2j)^d} \left(e^{-j\omega s} \sin\left(\frac{\omega}{2}\right)\right)^{d \bmod 2} \left(\sum_{m=0}^p \left(\cos\left(\frac{\omega}{2}\right)\right)^{2m}\right)^{\lfloor \frac{d+1}{2} \rfloor}, \quad (19)$$

where  $\lfloor x \rfloor$  denotes the largest integer smaller than  $x$ , the sampling shift for  $L(\omega)$  is the same as the one for  $H(\omega)$  (i.e.,  $s = \frac{(p+1) \bmod 2}{2}$ ), and the sampling shift for  $K(\omega)$  is the same as the one for  $G(\omega)$ .

Note that Equations (10) and (17)–(19) work fine from the mathematical point of view, but in practice the reconstruction may become cumbersome when both  $p$  and  $d$  are large. (The lengths of impulse responses  $h(n)$ ,  $g(n)$ ,  $l(n)$ , and  $k(n)$  are  $p+2$ ,  $d+1$ ,  $(p+1)(d-(d+1) \bmod 2)+1$ , and  $pd+(p+1)(d \bmod 2)+1$ , respectively, while for the frequency

responses of the decomposition filters we observe that  $\lim_{p \rightarrow \infty} |H_{-s}(\omega)| = \delta_u(\omega + 2n\pi)$  and  $\lim_{d \rightarrow \infty} (2j)^{-d} |G_{-s}(\omega)| = \delta_u(\omega + (2n+1)\pi)$  with  $n \in \mathbf{Z}$ .)

It is also worth noting that  $K(\omega)$  is a lowpass filter when  $p$  is even (i.e., the reconstruction function  $\chi(x)$  is a wavelet only for  $p$  odd).

Tables 1, 2, and 3 list impulse responses of the four filters for  $p \in \{0, 1, 2\}$  and  $d \in \{1, 2, 3\}$ , while Figure 4 shows wavelets  $\psi(x) = \frac{d^d \beta_{p+d}(x)}{dx^d}$  for the same values of  $p$  and  $d$ . Wavelets from this family have a support of length  $d+p+1$ , regularity order  $p$ , and are either symmetric or antisymmetric.

Table 1: Impulse responses  $h(n)$ ,  $g(n)$ ,  $l(n)$ , and  $k(n)$  for  $p=0$  and  $d \in \{1, 2, 3\}$ .

$n$	$h(n)$	$d = 1$			$d = 2$			$d = 3$		
		$g(n)$	$l(n)$	$k(n)$	$g(n)$	$l(n)$	$k(n)$	$g(n)$	$l(n)$	$k(n)$
-2								1		
-1	0.5	1			1			-3	-0.125	
0	0.5	-1	0.5	-0.25	-2	0.5	-0.25	3	0.625	0.0625
1			0.5	0.25	1	0.5		-1	0.625	-0.0625
2									-0.125	

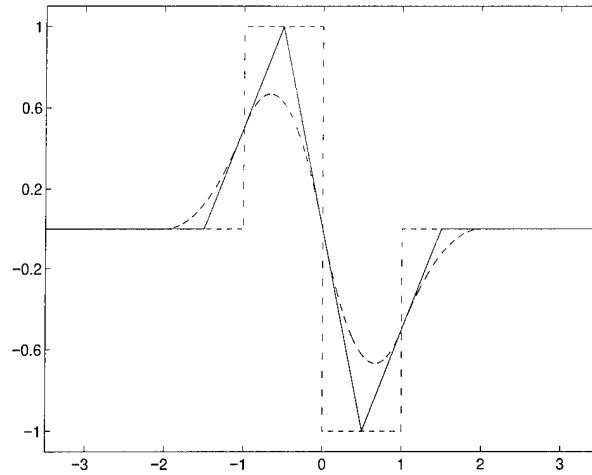
Table 2: Impulse responses  $h(n)$ ,  $g(n)$ ,  $l(n)$ , and  $k(n)$  for  $p=1$  and  $d \in \{1, 2, 3\}$ .

$n$	$h(n)$	$d = 1$		$d = 2$	
		$g(n)$	$k(n)$	$g(n)$	$k(n)$
-1	0.25	1	-0.0625	1	-0.0625
0	0.5	-1	-0.3125	-2	-0.375
1	0.25		0.3125	1	-0.0625
2			0.0625		

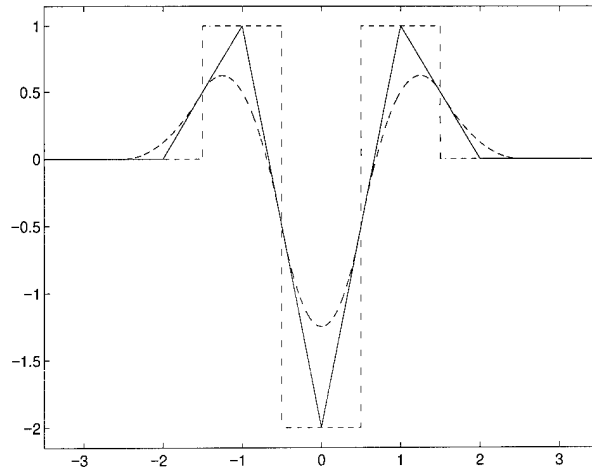
  

$n$	$d = 3$			
	$h(n)$	$g(n)$	$l(n)$	$k(n)$
-3			-0.015625	
-2		1	-0.09375	0.00390625
-1	0.25	-3	0.265625	0.04296875
0	0.5	3	0.6875	0.1015625
1	0.25	-1	0.265625	-0.1015625
2			-0.09375	-0.04296875
3			-0.015625	-0.00390625

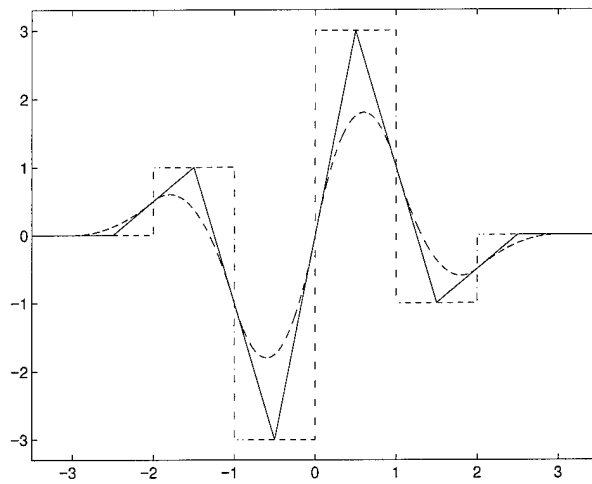
As already discussed, wavelets with  $p=2$  and  $d=1$  from a family of wavelets with  $p$  even and  $d=1$  were used in [14, 7], whereas filters with  $p=1$  and  $d=2$  from a family of filters with  $p$  odd and  $d=2$  were employed in [16, 17, 18]. Here described transform puts no restrictions on the choice of  $p$  or  $d$  whatsoever, and uses a better initialization procedure than the one originally proposed in [7].



(a)



(b)



(c)

Figure 4: (a) Wavelets (a)  $\psi(x) = \frac{d\beta_{p+1}(x)}{dx}$ , (b)  $\psi(x) = \frac{d^2\beta_{p+2}(x)}{dx^2}$ , and (c)  $\psi(x) = \frac{d^3\beta_{p+3}(x)}{dx^3}$  for  $p=0$  (dashdotted),  $p=1$  (solid), and  $p=2$  (dashed).

Table 3: Impulse responses  $h(n)$ ,  $g(n)$ ,  $l(n)$ , and  $k(n)$  for  $p=2$  and  $d \in \{1, 2, 3\}$ .

$n$	$h(n)$	$d = 1$			$d = 2$		
		$g(n)$	$l(n)$	$k(n)$	$g(n)$	$l(n)$	$k(n)$
-2	0.125			-0.015625			-0.015625
-1	0.375	1	0.125	-0.109375	1	0.125	-0.125
0	0.375	-1	0.375	-0.34375	-2	0.375	-0.46875
1	0.125		0.375	0.34375	1	0.375	-0.125
2			0.125	0.109375		0.125	-0.015625
3				0.015625			

$n$	$d = 3$			
	$h(n)$	$g(n)$	$l(n)$	$k(n)$
-4			-0.001953125	0.000244140625
-3			-0.017578125	0.003662109375
-2	0.125	1	-0.0703125	0.0263671875
-1	0.375	-3	0.085937	0.0908203125
0	0.375	3	0.50390625	0.13037109375
1	0.125	-1	0.50390625	-0.13037109375
2			0.0859375	-0.0908203125
3			-0.0703125	-0.0263671875
4			-0.017578125	-0.003662109375
5			-0.001953125	-0.000244140625

### 2.1.3 Steerable Functions

When extending the transform from Section 2.1.2 to two dimensions, one of the first questions that come to mind is how to deal with the fact that derivatives can be defined in any direction of the plane. In the case of a first derivative of a Gaussian, one can simply compute first derivatives of a Gaussian in  $x$  and  $y$  directions and then determine the derivative in any direction from these two directional derivatives [19]. For higher order derivatives of a Gaussian, Freeman and Adelson [20] showed that order+1 directional operators are needed for synthesizing the operator at any orientation. In fact, functions with the property of expressing their arbitrary rotations as linear combinations of some functions are not limited to derivatives of a Gaussian. Below, we briefly restate some of the results from [20].

A two-dimensional function is called “steerable” if its rotations generate a finite dimensional space. Rotations of a steerable function  $f(r, \theta)$  can therefore be expressed as

$$f(r, \theta - \theta_0) = \sum_{i=1}^I c_i(\theta_0) f_i(r, \theta), \quad (20)$$

where  $\theta_0$  denotes an arbitrary angle,  $\{c_i(\theta_0)\}$  stands for a set of interpolating functions,



$\{f_i(r, \theta)\}$  is a set of basis functions, and  $r = \sqrt{x^2 + y^2}$  and  $\theta = \arg(x, y)$  are polar radius and angle, respectively.

If  $f(r, \theta)$  represents a filter kernel, the result of filtering with a rotated filter  $f(r, \theta - \theta_0)$  can be computed simply by  $\{c_i(\theta_0)\}$  weighted linear combination of outputs from basis filters  $\{f_i(r, \theta)\}$ . Only the outputs from basis filters need to be precomputed and then the output from a filter rotated by any angle  $\theta_0$  can be found by interpolating between them. When a large number of rotations of a template filter is required, the above scheme can lead to substantial savings in both computational cost (time) and memory requirements (space). The necessary condition for a function  $f(r, \theta)$  to be steerable is that  $f(r, \theta)$  is bandlimited in its polar angle:

$$f(r, \theta) = \sum_{n=-N}^N a_n(r) e^{jn\theta}. \quad (21)$$

This can be verified by inserting (21) into (20) and by assuming, for convenience, that  $f_i(r, \theta) = f(r, \theta - \theta_i)$ :

$$a_n(r) e^{-jn\theta_0} = \sum_{i=1}^I c_i(\theta_0) a_n(r) e^{-jn\theta_i}, \quad (22)$$

where  $\{\theta_i\}$  is a set of user defined angles and  $n \in [-N, 0]$ .<sup>2</sup> Since only nonzero coefficients  $a_n(r)$  are of interest, both sides of (22) can be divided by  $a_n(r)$ . This yields a matrix equation from which interpolating functions  $c_i(\theta_0)$  can be determined:

$$\begin{bmatrix} 1 \\ e^{j\theta_0} \\ \vdots \\ e^{jN\theta_0} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \dots & e^{j\theta_I} \\ \vdots & \vdots & & \vdots \\ e^{jN\theta_1} & e^{jN\theta_2} & \dots & e^{jN\theta_I} \end{bmatrix} \begin{bmatrix} c_1(\theta_0) \\ c_2(\theta_0) \\ \vdots \\ c_I(\theta_0) \end{bmatrix}. \quad (23)$$

For coefficients  $a_n = 0$  the rows corresponding to each  $n$  were removed from the matrix formulation shown in (23). For this system to have a solution, the angles  $\{\theta_i\}$  must be chosen such that the columns of the matrix are linearly independent.

In [20] they proved that the minimum number of basis functions  $f_i(r, \theta)$  needed to steer  $f(r, \theta)$  according to (20) is equal to the number of nonzero coefficients  $a_n(r)$  in the Fourier series expansion (21).

To conclude this brief description of steerability, let us only remark that functions which are not steerable (i.e., do not have a finite number of terms in (21)) can be approximated with steerable functions (a singular value decomposition was employed for approximating

---

<sup>2</sup>Note that the constraints are the same for  $n \in [-N, -1]$  and  $n \in [1, N]$ , so that a subset of all possible values for  $n \in [-N, N]$  can be taken.

such functions efficiently [21]), and that the technique of expressing transformed versions of a function as linear combinations of a fixed set of basis functions is not limited to rotations (extensions to translations [22], scalings [21, 22], and general transformations [23] have been reported).

#### 2.1.4 Steerable Dyadic Wavelet Transform

Returning to the question from the beginning of Section 2.1.3, the answer seems obvious: one needs to construct a steerable analog to the one-dimensional transform from Section 2.1.2. Steerable transforms are nothing new—quite a few [24, 25, 26, 22] have been devised, some of them [25, 26] exactly for the computation of directional derivatives. Here, we are not interested in any directional derivatives: we want to use derivatives of central B-splines which, as the order of B-splines increases, tend to derivatives of a Gaussian.

We define a steerable dyadic wavelet transform of a function  $s(x, y) \in L^2(\mathbf{R}^2)$  at a scale  $2^m$ ,  $m \in \mathbf{Z}$ , as [27]

$$W_m^i s(x, y) = s * \psi_m^i(x, y), \quad (24)$$

where  $\psi_m^i(x, y)$  denotes  $\psi_m(x, y)$  rotated by  $\theta_i$ ,  $\psi_m(x, y) = 2^{-2m}\psi(2^{-m}x, 2^{-m}y)$ ,  $\psi(x, y)$  is a steerable wavelet that can be steered with I basis functions, and  $\theta_i = \frac{i-1}{I}\pi$  with  $i \in \{1, 2, \dots, I\}$ .

Analogously to the one-dimensional case (cf. Section 2.1.2) we require the two-dimensional Fourier plane to be covered by the dyadic dilations of  $\hat{\psi}^i(2^m\omega_x, 2^m\omega_y)$ : there must exist  $A_3 > 0$  and  $B_3 < \infty$  such that

$$A_3 \leq \sum_{m=-\infty}^{\infty} \sum_{i=1}^I |\hat{\psi}^i(2^m\omega_x, 2^m\omega_y)|^2 \leq B_3 \quad (25)$$

is satisfied almost everywhere.

If (nonunique) reconstructing functions  $\chi_m^i(x, y)$  are chosen such that their Fourier transforms satisfy

$$\sum_{m=-\infty}^{\infty} \sum_{i=1}^I \hat{\psi}^i(2^m\omega_x, 2^m\omega_y) \hat{\chi}^i(2^m\omega_x, 2^m\omega_y) = 1, \quad (26)$$

the function  $s(x, y)$  may be reconstructed from its steerable dyadic wavelet transform by

$$s(x, y) = \sum_{m=-\infty}^{\infty} \sum_{i=1}^I W_m^i s * \chi_m^i(x, y), \quad (27)$$

where  $\chi_m^i(x, y)$  denotes  $\chi_m(x, y)$  rotated by  $\theta_i$  and  $\chi_m(x, y) = 2^{-2m}\chi(2^{-m}x, 2^{-m}y)$ .

To derive an algorithm for the fast computation of the transform, we, similar to (3), introduce two smoothing functions such that

$$\hat{\phi}(\omega_x, \omega_y) \hat{\varphi}(\omega_x, \omega_y) = \sum_{m=0}^{\infty} \sum_{i=1}^I \hat{\psi}^i(2^m \omega_x, 2^m \omega_y) \hat{\chi}^i(2^m \omega_x, 2^m \omega_y). \quad (28)$$

We choose wavelets that are steerable analogs to the one-dimensional wavelets from Section 2.1.2:<sup>3</sup>

$$\hat{\psi}(\omega_r, \omega_\theta) = (j\omega_r \cos(\omega_\theta))^d \left( \frac{\sin(\frac{\omega_r}{2})}{\frac{\omega_r}{2}} \right)^{p+d+1}, \quad (29)$$

where  $\omega_r = \sqrt{\omega_x^2 + \omega_y^2}$  and  $\omega_\theta = \arg(\omega_x, \omega_y)$ . These wavelets can be steered with  $d+1$  basis functions (i.e.,  $I$  in (20) is equal to  $d+1$ ).

Choosing

$$\hat{\phi}(2\omega_r) = H_{st}(\omega_r) \hat{\phi}(\omega_r), \quad (30)$$

$$\hat{\psi}^i(\omega_r, \omega_\theta) = G_{st}(\omega_r, \omega_\theta - \theta_i) \hat{\phi}(\omega_r), \quad (31)$$

$$\hat{\varphi}(2\omega_r) = L_{st}(\omega_r) \hat{\varphi}(\omega_r), \quad (32)$$

and

$$\hat{\chi}^i(\omega_r, \omega_\theta) = K_{st}(\omega_r, \omega_\theta - \theta_i) \hat{\varphi}(\omega_r), \quad (33)$$

and using (30) through (33) with (28) computed for the finest two scales, we obtain

$$\sum_{i=1}^I G_{st}(\omega_r, \omega_\theta - \theta_i) K_{st}(\omega_r, \omega_\theta - \theta_i) + H_{st}(\omega_r) L_{st}(\omega_r) = 1. \quad (34)$$

Setting  $\hat{\phi}(\omega_r) = \hat{\beta}_p(\omega_r)$ , and employing (6) and (29) with (30) and (31), we find that

$$H_{st}(\omega_r) = H_{-s}(\omega_r) \quad (35)$$

and

$$G_{st}(\omega_r, \omega_\theta) = (\cos(\omega_\theta))^d G_{-s}(\omega_r), \quad (36)$$

where  $H(\omega)$  and  $G(\omega)$  are specified by (10) and (17), respectively.

By inserting (35) and (36) into (34), the missing two filters can be chosen as

$$L_{st}(\omega_r) = L_s(\omega_r) \quad (37)$$

and

$$K_{st}(\omega_r, \omega_\theta) = \frac{1}{C_d} (\cos(\omega_\theta))^d K_s(\omega_r), \quad (38)$$

where  $L(\omega)$  and  $K(\omega)$  are given by (18) and (19), respectively, and  $C_d = \sum_{i=1}^I (\cos(\omega - \theta_i))^{2d}$ .

---

<sup>3</sup>This choice can be viewed as an extension of the transform presented in [28, 27, 29].

### 2.1.5 Multiscale Spline Derivative-Based Transform

Let us pause here for a brief assessment of the two-dimensional steerable transform derived so far. We have chosen steerable wavelets (29) which are equal to  $d$ -th order derivatives of circularly symmetric spline functions in the direction of  $x$ -axis (note that knots for these splines are circles) and we have laid a foundation for filter bank implementations in (34). An obvious shortcoming of this scheme is the fact that none of the filter kernels obtained from (35) through (38) is compactly supported on the rectangular grid. For implementations using digital filters, one is therefore forced to approximate these frequency responses, and by doing so, (34) may not hold anymore. Filters in filter bank implementations of steerable pyramids described in [25, 26, 22], for example, were designed by using various techniques for approximating desired frequency responses. None of the reported filter banks achieved perfect reconstruction and all filter kernels were nonseparable. Here, we will take a different approach. We will approximate the wavelets in (29) in a way that will lead to  $x$ - $y$  separable filters in a perfect reconstruction filter bank implementation of the transform such that the quality of approximation will improve by increasing the order of B-splines.

Let us approximate wavelets from (29) with

$$\psi(x, y) = \frac{d^d \beta_{p+d}(x)}{dx^d} \beta_{p+d}(y). \quad (39)$$

Based on the fact that B-splines tend to a Gaussian as their order increases, it is easy to see that both wavelets (29) and (39) converge to the same functions (i.e.,  $d$ -th order derivatives of the normalized Gaussian in the direction of  $x$ -axis) as  $p \rightarrow \infty$ .

In order to steer wavelets  $\psi(x, y)$  given by (39) (note that steering will be only approximate, since these wavelets are not steerable), we need to find basis functions that will approximately steer  $\psi(x, y)$ . Computing rotations, as we did in (24), is not practical here, because arbitrary rotations of (39) cannot be expressed exactly in terms of central B-spline functions. Instead, we take advantage of the property of circularly symmetric functions that rotations of their directional derivatives are equal to directional derivatives in rotated directions:

$$\mathcal{R}_{\theta_0} \left\{ \frac{\partial^d \varrho_c(x, y)}{\partial \vec{n}^d} \right\} = \frac{\partial^d \varrho_c(x, y)}{\partial \vec{n}_{\theta_0}^d},$$

where  $\mathcal{R}_{\theta_0}$  stands for rotation by  $\theta_0$ ,  $\frac{\partial \varrho_c(x, y)}{\partial \vec{n}} = \vec{n} \cdot \vec{\nabla} \varrho_c(x, y)$ ,  $\varrho_c(x, y)$  is a circularly symmetric function, and  $\vec{n}_{\theta_0}$  denotes vector  $\vec{n} = (\cos \theta, \sin \theta)$  rotated by  $\theta_0$ .

Let us choose

$$\varrho(x, y) = \beta_{p+d}(x) \beta_{p+d}(y),$$

which is approximately circularly symmetric function for higher order splines. A rotation of  $\psi(x, y) = \frac{\partial^d \varrho(x, y)}{\partial x^d}$  from (39) by  $\theta_0$  can therefore be approximated by

$$\psi^{\theta_0}(x, y) \simeq \frac{\partial^d \varrho(x, y)}{\partial \vec{n}^d} = \sum_{i=0}^d \binom{d}{i} n_x^{d-i} n_y^i \frac{d^{d-i} \beta_{p+d}(x)}{dx^{d-i}} \frac{d^i \beta_{p+d}(y)}{dy^i}, \quad (40)$$

where  $\vec{n} = (\cos \theta_0, \sin \theta_0) = (n_x, n_y)$ .

Note that in case of Gaussian, which is both  $x$ - $y$  separable and circularly symmetric, (40) becomes exact (e.g., for  $\varrho(x, y) = e^{-(x^2+y^2)}$ ,  $\theta_0 = -\theta$ , and  $d = \{2, 4\}$ , we obtain, up to a scaling factor,  $x$ - $y$  separable basis set for the second and fourth derivative of Gaussian from Tables III and VII of [20]).

Having found a set of basis functions (40) that approximately steer wavelets (39), we want to construct a transform such that Equations (24) through (28) will be valid (superscript  $i$  must be viewed now as an index, rather than rotation by  $\theta_i$ ). In frequency domain, we can express basis functions from (40) as

$$\hat{\psi}^{i+1}(\omega_x, \omega_y) = G_{-s}^{d-i}(\omega_x) G_{-s}^i(\omega_y) \hat{\beta}_{p+i}(\omega_x) \hat{\beta}_{p+d-i}(\omega_y), \quad i = 0, 1, \dots, d, \quad (41)$$

where  $G^d(\omega)$  is given by (17) and  $G^0(\omega) = 1$ .

Choosing appropriate  $\hat{\chi}^i(\omega_x, \omega_y)$  to obtain a relation needed for the filter bank implementation of the transform is more complicated than in one dimension. Since we could not find a general solution for arbitrary  $d$ , we solve each case separately. Below, we present solutions for the first four orders. When  $d \leq 2$ , we impose  $\hat{\varphi}(\omega_x, \omega_y) = \hat{\phi}(\omega_x, \omega_y) = \hat{\beta}_p(\omega_x) \hat{\beta}_p(\omega_y)$ , a constraint analogous to the one from Section 2.1.2. For  $d = 1$ , we write similar to [7]

$$\hat{\chi}^1(\omega_x, \omega_y) = K_s^1(\omega_x) T_1(\omega_y) \hat{\beta}_p(\omega_x) \hat{\beta}_{p-1}(\omega_y), \quad (42)$$

$$\hat{\chi}^2(\omega_x, \omega_y) = T_1(\omega_x) K_s^1(\omega_y) \hat{\beta}_{p-1}(\omega_x) \hat{\beta}_p(\omega_y), \quad (43)$$

where

$$T_1(\omega) = \frac{1}{2} (1 + |H(\omega)|^2)$$

and  $K^d(\omega)$  is given by (19).

Computing (28) for the finest two scales and inserting (6), (41), (42), and (43) yields a relation between frequency responses

$$G^1(\omega_x) K^1(\omega_x) T_1(\omega_y) + T_1(\omega_x) G^1(\omega_y) K^1(\omega_y) + |H(\omega_x) H(\omega_y)|^2 = 1.$$

For  $d = 2$ , we choose

$$\hat{\chi}^1(\omega_x, \omega_y) = K_s^2(\omega_x)T_2(\omega_y)\hat{\beta}_p(\omega_x)\hat{\beta}_{p-2}(\omega_y), \quad (44)$$

$$\hat{\chi}^2(\omega_x, \omega_y) = K_s^1(\omega_x)K_s^1(\omega_y)\hat{\beta}_{p-1}(\omega_x)\hat{\beta}_{p-1}(\omega_y), \quad (45)$$

$$\hat{\chi}^3(\omega_x, \omega_y) = T_2(\omega_x)K_s^2(\omega_y)\hat{\beta}_{p-2}(\omega_x)\hat{\beta}_p(\omega_y), \quad (46)$$

where

$$T_2(\omega) = |H(\omega)|^2. \quad (47)$$

Using (6), (41), and (44) through (46) with (28) results in

$$G^2(\omega_x)K^2(\omega_x)T_2(\omega_y) + G^1(\omega_x)K^1(\omega_x)G^1(\omega_y)K^1(\omega_y) + T_2(\omega_x)G^2(\omega_y)K^2(\omega_y) + |H(\omega_x)H(\omega_y)|^2 = 1.$$

For orders  $d > 2$ , we require  $\hat{\phi}(\omega_x, \omega_y) = \hat{\beta}_p(\omega_x)\hat{\beta}_p(\omega_y)$  and  $\hat{\varphi}(\omega_x, \omega_y) = \hat{\varphi}(\omega_x)\hat{\varphi}(\omega_y)$ , where  $\hat{\varphi}(\omega)$  is specified by (8) and (18).

For  $d = 3$ , we choose reconstructing functions

$$\hat{\chi}^1(\omega_x, \omega_y) = K_s^3(\omega_x)\hat{\varphi}(\omega_x)\hat{\varphi}_{-3}(\omega_y), \quad (48)$$

$$\hat{\chi}^2(\omega_x, \omega_y) = -K_s^2(\omega_x)K_s^1(\omega_y)V_3(\omega_x)V_3(\omega_y)\hat{\varphi}_{-1}(\omega_x)\hat{\varphi}_{-2}(\omega_y), \quad (49)$$

$$\hat{\chi}^3(\omega_x, \omega_y) = -K_s^1(\omega_x)K_s^2(\omega_y)V_3(\omega_x)V_3(\omega_y)\hat{\varphi}_{-2}(\omega_x)\hat{\varphi}_{-1}(\omega_y), \quad (50)$$

$$\hat{\chi}^4(\omega_x, \omega_y) = K_s^3(\omega_y)\hat{\varphi}_{-3}(\omega_x)\hat{\varphi}(\omega_y), \quad (51)$$

where

$$V_3(\omega) = \frac{1}{\sqrt{2}}(1 - |H(\omega)|^2), \quad (52)$$

and  $\hat{\varphi}_{-i}(\omega) \in L^1(\mathbf{R})$  denotes a function such that  $\hat{\varphi}(\omega) = \hat{\beta}_{i-1}(\omega)\hat{\varphi}_{-i}(\omega)$ ,  $i \in \mathbf{N}$ .

Employing (41), (6), (8), and (48) through (51) with (28) yields a relation

$$G^3(\omega_x)K^3(\omega_x) - G^2(\omega_x)K^2(\omega_x)V_3(\omega_x)G^1(\omega_y)K^1(\omega_y)V_3(\omega_y) - G^1(\omega_x)K^1(\omega_x)V_3(\omega_x)G^2(\omega_y)K^2(\omega_y)V_3(\omega_y) + G^3(\omega_y)K^3(\omega_y) + H(\omega_x)L(\omega_x)H(\omega_y)L(\omega_y) = 1,$$

where  $L(\omega)$  is specified by (18).

For  $d = 4$ , our choices are

$$\hat{\chi}^1(\omega_x, \omega_y) = K_s^4(\omega_x)T_2(\omega_y)\hat{\varphi}(\omega_x)\hat{\varphi}_{-4}(\omega_y), \quad (53)$$

$$\hat{\chi}^2(\omega_x, \omega_y) = K_s^3(\omega_x)K_s^1(\omega_y)\hat{\varphi}_{-1}(\omega_x)\hat{\varphi}_{-3}(\omega_y), \quad (54)$$

$$\hat{\chi}^3(\omega_x, \omega_y) = -K_s^2(\omega_x)K_s^2(\omega_y)V_4(\omega_x)V_4(\omega_y)\hat{\varphi}_{-2}(\omega_x)\hat{\varphi}_{-2}(\omega_y), \quad (55)$$

$$\hat{\chi}^4(\omega_x, \omega_y) = K_s^1(\omega_x)K_s^3(\omega_y)\hat{\varphi}_{-3}(\omega_x)\hat{\varphi}_{-1}(\omega_y), \quad (56)$$

$$\hat{\chi}^5(\omega_x, \omega_y) = T_2(\omega_x)K_s^4(\omega_y)\hat{\varphi}_{-4}(\omega_x)\hat{\varphi}(\omega_y), \quad (57)$$

where

$$V_4(\omega_x) = 1 - |H(\omega)|^2. \quad (58)$$

Using the above functions (53) through (57), (41), (6), and (8) in (28) computed for the finest two scales gives

$$\begin{aligned} & G^4(\omega_x)K^4(\omega_x)T_2(\omega_y) + G^3(\omega_x)K^3(\omega_x)G^1(\omega_y)K^1(\omega_y) - \\ & - G^2(\omega_x)K^2(\omega_x)V_4(\omega_x)G^2(\omega_y)K^2(\omega_y)V_4(\omega_y) + G^1(\omega_x)K^1(\omega_x)G^3(\omega_y)K^3(\omega_y) + \\ & + T_2(\omega_x)G^4(\omega_y)K^4(\omega_y) + H(\omega_x)L(\omega_x)H(\omega_y)L(\omega_y) = 1. \end{aligned}$$

Here, we have even more freedom for choosing the reconstructing functions than in one dimension. The above functions for  $d = \{2, 3, 4\}$  were found by trying to reproduce the one-dimensional transform from Section 2.1.2 as much as possible. All decomposition filters  $G^d(\omega)$  were first paired with corresponding reconstruction filters  $K^d(\omega)$ , and then all other potential digital filters were specified as polynomials in  $H_{-s}(\omega)$ . We inserted thus specified filters into a relation between their frequency responses and solved for the unknown polynomial coefficients. Since we allowed more filters with higher-degree polynomials than expected in the solution, the resulting system of linear equations was underdetermined. This allowed enough freedom for removal of undesired digital filters and for balance between degrees of polynomials.

The described procedure for determination of reconstructing functions and filters involves quite a lot of heuristics to obtain the appropriate solution from the underdetermined linear system. Unfortunately, we are not aware of any systematic way (aside from numerical optimization, which may be pretty cumbersome) to obtain solutions comparable to the ones above.

Next, we will derive a filter bank implementation of the transform. Similar to the one-dimensional case from Section 2.1.2, we assume a bandlimited input signal:

$\hat{s}(\omega_x, \omega_y) = 0$  for  $|\omega_x| > \pi$  or  $|\omega_y| > \pi$ . Using Shannon's sampling theorem in two dimensions [30] with (24) and basis functions from (41), we can write

$$\begin{aligned} W_0^i s(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{i_x=-\infty}^{\infty} \sum_{i_y=-\infty}^{\infty} s(i_x, i_y) \text{sinc}(t_x - i_x) \text{sinc}(t_y - i_y) \cdot \\ &\cdot \sum_{m_x=-\infty}^{\infty} g_{-s}^{d-i}(m_x) \beta_{p+i}(x - t_x - m_x) \sum_{m_y=-\infty}^{\infty} g_{-s}^i(m_y) \beta_{p+d-i}(y - t_y - m_y) dt_x dt_y, \end{aligned}$$

where  $i = 0, 1, \dots, d$  as in (41).

Again, we approximate sinc functions with  $r$ -order cardinal splines, then use (12), and get

$$\begin{aligned} \text{DFT}\{W_m^i s(x, y)|_{x=n_x, y=n_y}\} &\simeq S(\omega_x, \omega_y) B_r^{-1}(\omega_x) B_r^{-1}(\omega_y) B_{p+r+i+1}(\omega_x) \cdot \\ &\cdot B_{p+r+d-i+1}(\omega_y) G_{-s}^{d-i}(2^m \omega_x) G_{-s}^i(2^m \omega_y) \prod_{n=0}^{m-1} H_{-s}^{p+i}(2^n \omega_x) H_{-s}^{p+d-i}(2^n \omega_y). \end{aligned} \quad (59)$$

Using (59) with an approximation  $B_{p+r+i+1}(\omega) \simeq B_{p+r}(\omega) B_i(\omega)$ , we can obtain a filter bank implementation of the transform decomposition. The reconstruction part follows from (28), (41), and reconstructing functions for distinct values of  $d$ . Figure 5 shows filter bank implementations of a multiscale spline derivative-based transform for  $d = \{1, 2, 3, 4\}$ . For  $d = 1$ , we recognize (except for the prefiltering and postfiltering) the filter bank implementation of a two-dimensional discrete dyadic wavelet transform from [7]. For  $d = 2$ , however, our transform differs from the filter bank presented in [17]: second derivative is computed only in the directions of  $x$  and  $y$ -axis in [31, 17], which is not enough for steering. Although not particularly appropriate for orientation analysis, such a scheme may still efficiently approximate Laplacian of Gaussian across dyadic scales.

A transform similar to the one described in this section, was presented in [25, 26, 22].

Their filter bank implementation is less redundant (downsampling is used on the lowpass branch, while simultaneously keeping aliasing negligible by using a filter with insignificant amount of energy for  $|\omega_r| > \frac{\pi}{2}$ ) and uses reconstruction filters with same magnitude frequency responses as the decomposition ones—a possible advantage for certain applications. They have, on the other hand, little control over the function from which derivatives are computed (to obtain a  $d$ -th derivative, they multiply a circularly symmetric filter by  $(-j \cos \theta)^d$  with all filters being obtained by recursive minimization of a weighted combination of constraints), filter bank does not have perfect reconstruction, and none of the filters is  $x$ - $y$  separable.

### 2.1.6 Image Fusion Application

A new direction in our visualization methodology (Phase IV statement of work) involves the use of image fusion. A problem with our early algorithms using multiscale global enhancement was that they produced results which were somewhat unfamiliar looking compared to traditional methods of enhancement and the original mammogram. Examples of two such cases are shown in the middle of Figures 6(a) and 6(b): textures of fibrous tissues extracted from the modulus of transform coefficients and multiscale histogram equalization [32]. These representations did an outstanding job of extracting salient features in mammograms but were difficult for radiologists to interpret without additional training.



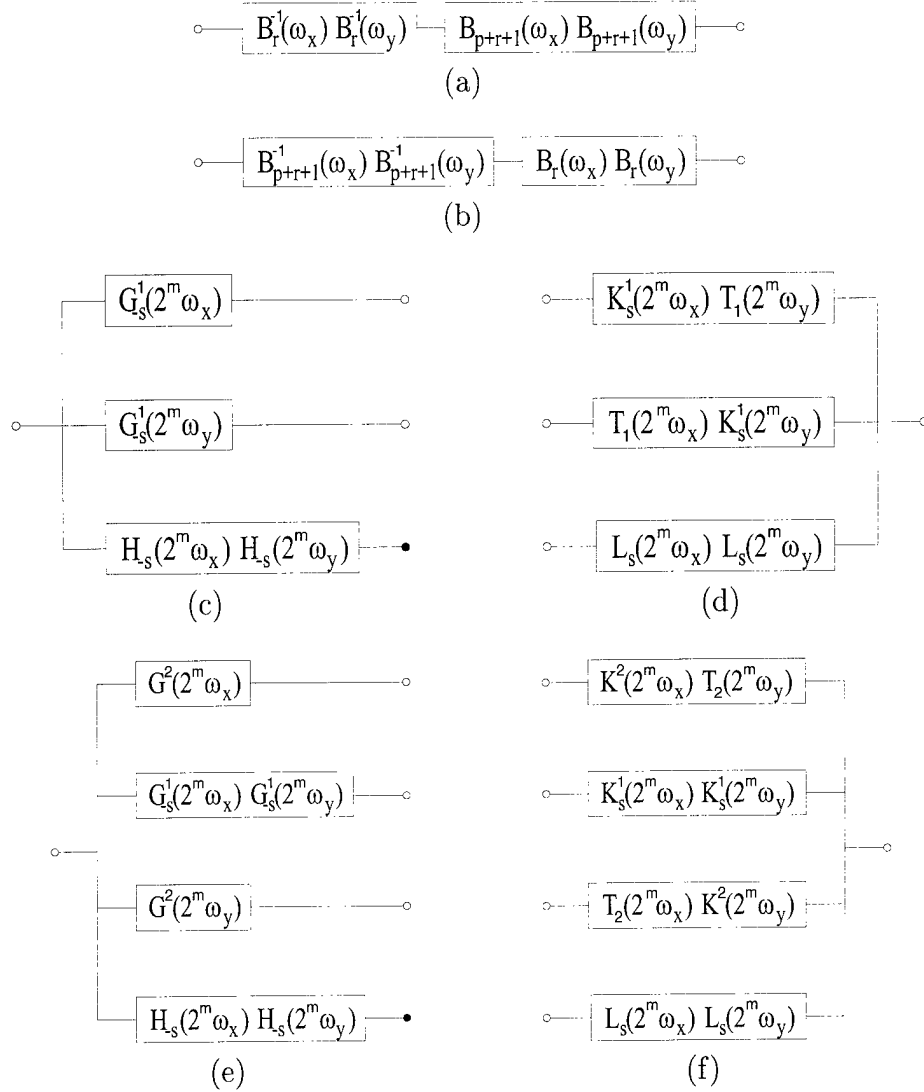


Figure 5: Filter bank implementation of a multiscale spline derivative-based transform for  $m \in [0, M - 1]$ : (a) Prefiltering, (b) postfiltering, (c) decomposition and (d) reconstruction modules for  $d = 1$ , and (e) decomposition and (f) reconstruction modules for  $d = 2$ .

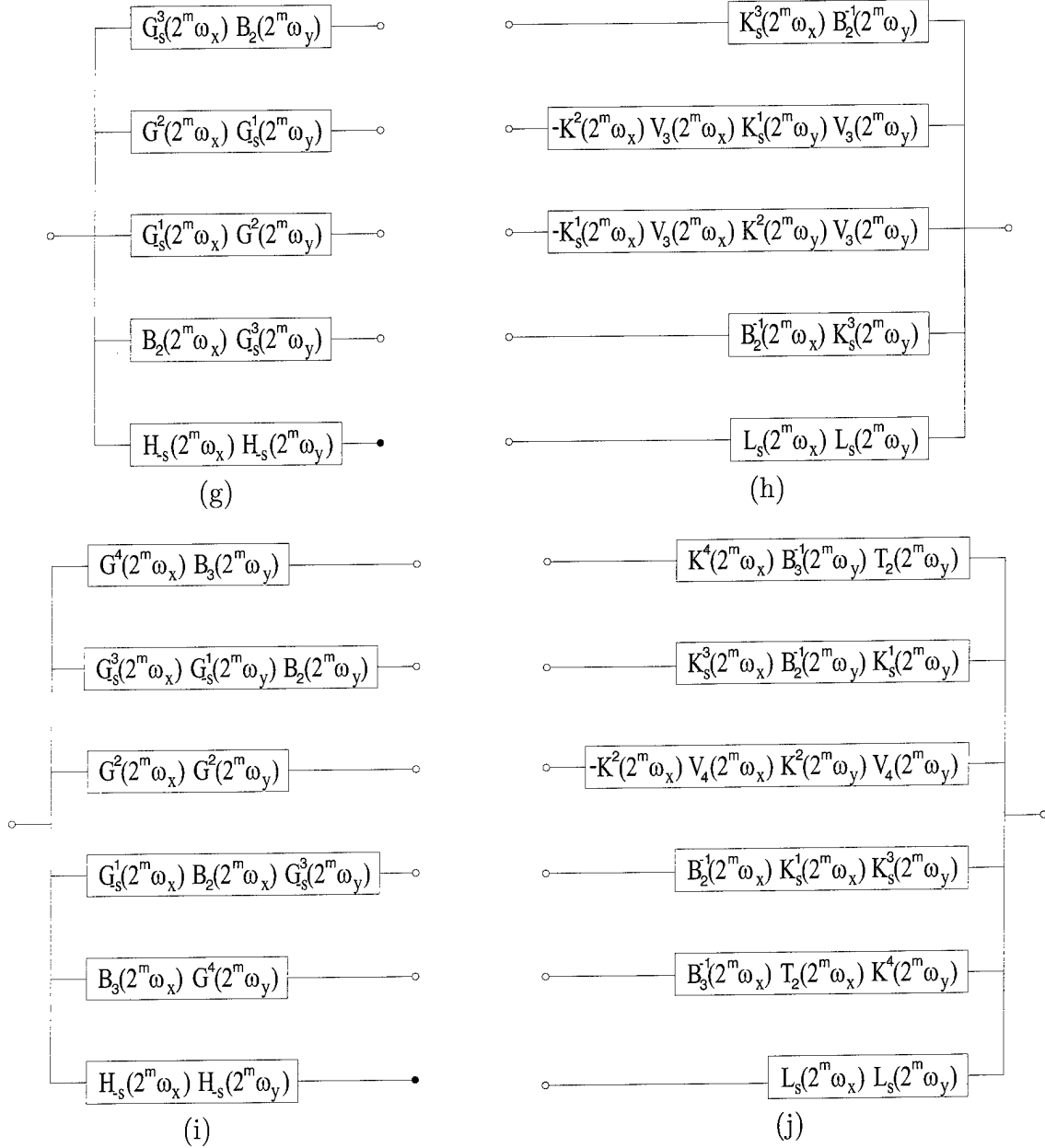


Figure 5: Continued: (g) Decomposition and (h) reconstruction modules for  $d = 3$ , and (i) decomposition and (j) reconstruction modules for  $d = 4$ . Decomposition modules are recursively applied at the locations of the filled circles.

By integrating the original mammogram with (one or more) enhanced versions we can produce (synthesize) an image which is more familiar (and appealing) to mammographers and general radiologists. Thus, fusion plays an important role in avoiding a “paradigm shift” in current protocols.

Traditionally, image fusion combines different aspects of information from the same imaging modality or from distinct imaging modalities [33] and can be used to improve the reliability of a particular computational vision task or to provide a human observer with a deeper insight about the nature of observed data.

Our fusion method is based upon a multiscale spline derivative-based transform [28] which enables multiscale image processing along arbitrary orientations. Steerable dyadic wavelet transforms share many properties with multiscale representations used in image fusion, such as pyramids [34, 35] and traditional wavelet analysis [36, 37]. However an important difference is translation and rotation invariance, and the absence of aliasing in its filter bank implementation. Both aliasing and translation non-invariance can be sources of unwanted artifacts in a fused image [38].

For images to be fused, filters from the filter bank implementation of a multiscale spline derivative-based transform are used in quadrature pairs (i.e., with their Hilbert transform counterparts). Quadrature pairs of filters steered to some arbitrary angle are used to determine the local oriented energy defined as the sum of the squared outputs from each filter of the quadrature pair. Local dominant orientation (i.e., the angle that maximizes the local oriented energy) is determined at each level and position, filters are steered to the local dominant orientation, and local oriented energies are compared. The coefficients corresponding to the greater local oriented energy are included for reconstruction, resulting in a fused image.

Figure 6(a) shows an example of fusion of a mammographic image with the modulus of the dyadic wavelet coefficients, while Figure 6(b) presents an example of fusion of a mammographic image with the mammogram enhanced by multiscale histogram equalization. In both cases the fused image features enhanced structures, while having an appearance similar to the original mammogram.

### 2.1.7 Summary

The one-dimensional discrete dyadic wavelet transform was extended to higher-order derivatives and even-order spline functions, and an improved initialization procedure was developed. In comparison to the originally employed initialization [7], our method showed significantly better performance at finer scales of analysis (of importance for mammographic features including microcalcifications).

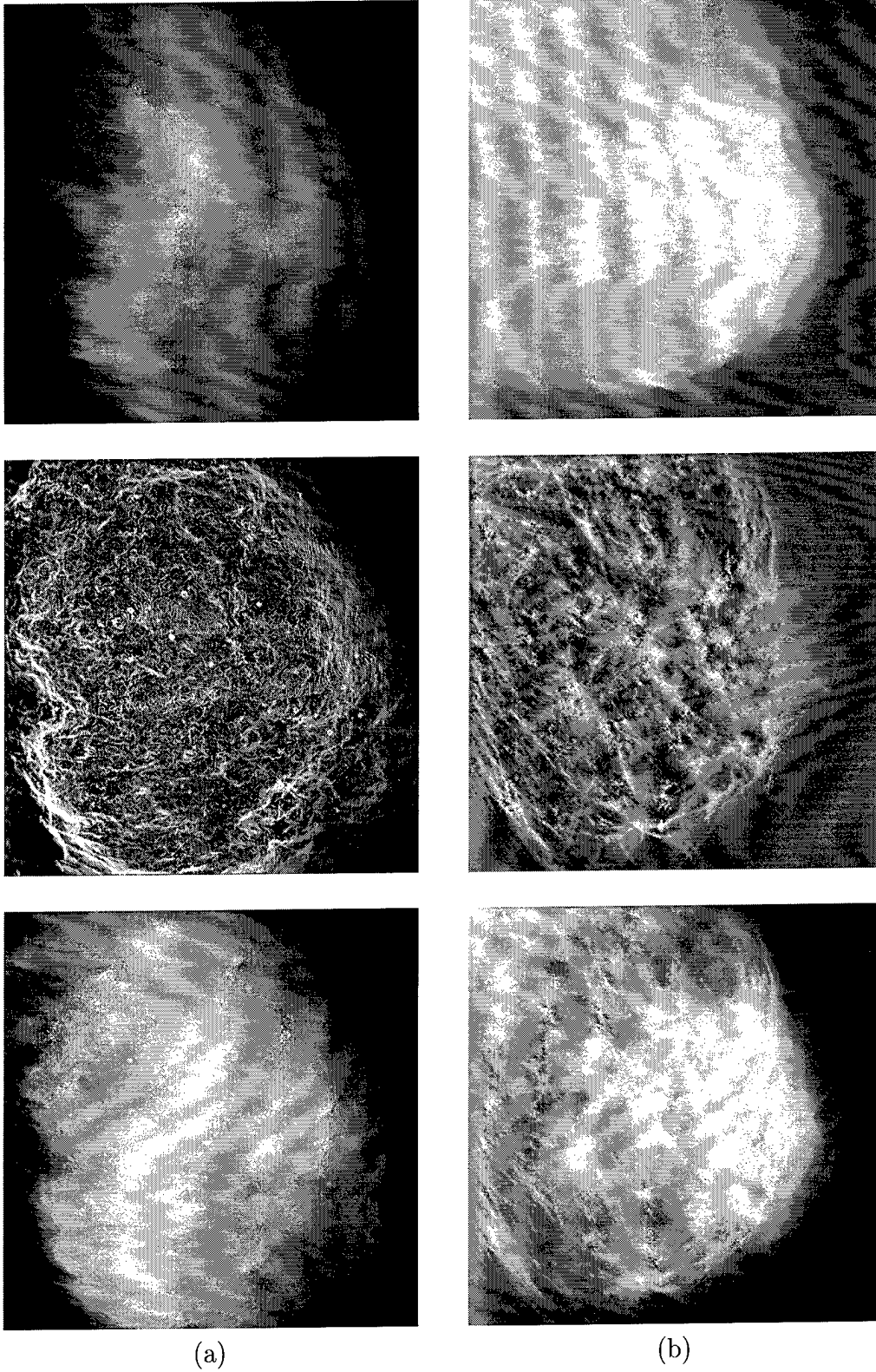


Figure 6: (a): Fused image (bottom) of an original mammogram (top) with the selected modulus of dyadic wavelet coefficients (middle). (b) Fusion (bottom) of a dense mammogram (top) with a mammogram enhanced by multiscale histogram equalization (middle).

A multiscale spline derivative-based transform was constructed as an approximation to a steerable dyadic wavelet transform. The transform used  $x$ - $y$  separable filters in a perfect reconstruction filter bank and enabled fast directional analysis of images without the introduction of artifacts due to translation and rotation invariance. Such artifacts are inherent to traditional methods of wavelet analysis.

Preliminary image fusion results showed promise for producing enhanced images with a more familiar appearance to radiologists.

## **2.2 A Parallel Algorithm to Support Interactive Wavelet Processing on a Radiologist Workstation**

### **2.2.1 Introduction**

Diagnosis of mammography is one of medical imaging applications where digital image processing is used to extract either implicit or explicit characteristics from digital data. Many of medical imaging applications require broad base-level functionality for image manipulation, and display to support technical end-users in areas of Radiology and CAD (Computer Aided Diagnosis). An electronic view-box for medical imaging analysis may require advanced image processing capabilities, including transformation, filtering or enhancement [39]. In addition, common image manipulation facilities such as convolution, warping, logical operations, comparisons, and arithmetic are often desirable to exploit the large dynamic range in contrast resolution provided by digital X-ray detectors.

In this chapter we describe the design of a high-speed parallel algorithm to support interactive analysis via multiscale processing. In addition to writing multi-threaded code, based on kernel support for multiple threads of control and changing the scheduling class policy to real-time mode, our algorithm required hardware and software support for high-speed image manipulation.

We used a SPARCstation 20 Model 71 with an SX frame buffer to execute our algorithm. The SPARCstation SX system's direct accessibility to physical and display memory provided flexible and powerful capabilities for the management, manipulation, and display of large image matrix sizes, typical of digital mammography

On top of the SPARCstation SX hardware, foundation graphics libraries were utilized. XGL and XIL accelerate many of graphics and X-extension library functions. As a foundation-level imaging interface for the SUN Unix programming environment, the XIL library defines how imaging operations such as display and image manipulation are carried out on the workstation [40].

In this section, we review the architecture of the SPARCstation SX system, and

describe briefly the XIL foundation graphics interface and relevant functions accelerated on the SX system processor. We then introduce new algorithm to accomplish the forward and inverse discrete dyadic wavelet transforms that exploits the parallel architecture of the machine.

## 2.2.2 Architectural Overview of SPARCstation SX System

The SPARCstation SX graphics processor is a scalable graphics architecture consisting of an enhancement to memory controller chip on an MBus within a SPARCstation. The SPARCstation SX system relies on the notion that graphics operations become extensions of native integer and floating-point operations of the CPU. The SPARCstation SX system employs a closely-coupled architecture, wherein a dedicated processor handles low-level display operations through processor-memory direct access. Residing directly on the CPU-memory bus, the SPARCstation SX memory controller, called the SMC, can perform extremely fast transfer of data from physical memory (DRAM) to the display frame buffer. The close-coupling of the SMC to the SPARCstation CPU further enables floating point-based computation to utilize the high-performance transformation capabilities on the SPARC processor. Figure 7 illustrates the high-level architecture of the SPARCstation MBus system, including CPU, SMC processor, and video memory.

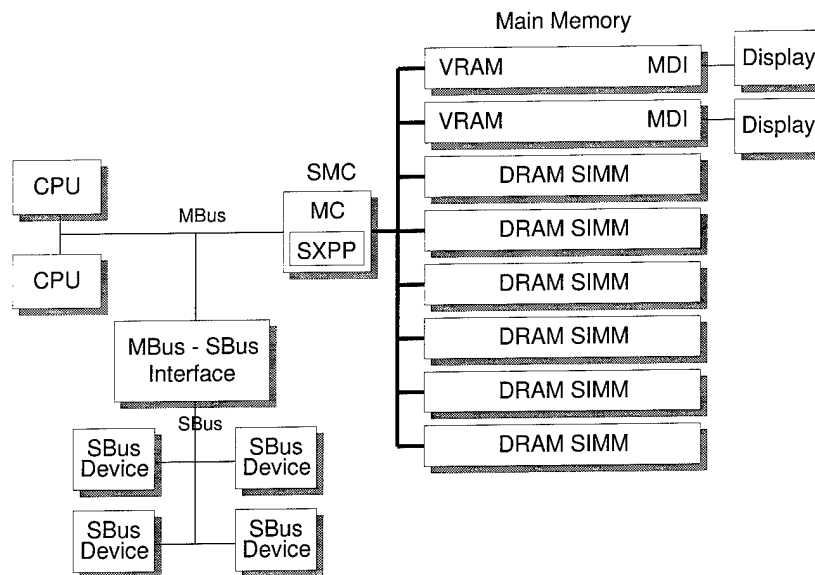


Figure 7: SPARCstation SX system high-level architecture [39].

The SMC processor lies on the MBus between the SPARC CPU(s) and memory. In addition, Video RAM (VRAM) is mapped into the same address space, allowing the SMC

to access and manipulate the display in the same manner as DRAM. This addressing gives the SPARCstation SX system superior bandwidth in moving large size images, e.g. mammograms, from memory to the screen.

Instructions executing in one or more of the SPARC CPUs make requests of the memory controller (MC) portion of the processor. The memory controller logic fulfills CPU requests for memory accesses, such as those required for fetching, loading and storing instructions.

When one of the CPUs encounters a graphics operation, it passes it to the SMC pixel processor (SXPP), which executes the instruction directly. When memory access is required to complete an operation, the SXPP makes a request to the memory controller.

### 2.2.3 XIL Imaging Library

The XIL library from Sun Microsystems is a foundation-level imaging interface providing common functions required by most imaging applications. Such a library is an application programming interface (API), but it is different from “normal” APIs because it is hardware dependent. The XIL library contains three primary components: a programming interface specification for basic imaging functionality, a high-performance implementation of the specification, and a standard hardware interface specification, which enables third-party hardware developers to readily support the XIL library and applications built on it [40]. Figure 8 illustrates the XIL library in relationship to other foundation-level interfaces.

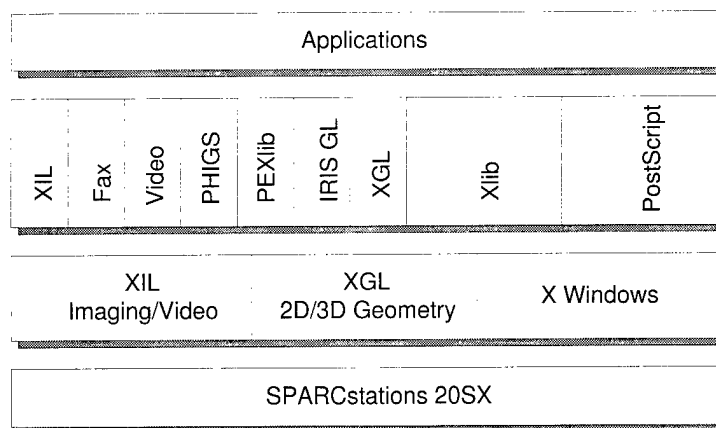


Figure 8: Solaris foundation graphics libraries and layered interfaces [40].

*Monadic* and *dyadic* operations are of particular interest within XIL library. Every XIL operator can be considered an atom. Groups of XIL operators concatenated in a sequential manner are known as molecules. Molecules are important to the SPARCstation

SX SMC processor because they allow the device to notify the XIL library at runtime, and obtain hardware support for groups of chained operations. At runtime, the SMC device handler loads into memory and notifies the XIL library of the atoms and molecules it wishes to replace with its own equivalent mapping to hardware. The XIL library uses a directed acyclic graph to hold a list of these dependencies during execution. When the XIL library encounters a known sequence of operations, it transfers the operation to the SMC for execution. XIL applications can realize significant performance improvement from the deferred execution [40].

## 2.2.4 An Algorithm for High-Speed Wavelet Analysis

In frame-based multi-scale representations [32], spatial convolution require  $2^l - 1$  zeros between each non-zero filter coefficients in a filter kernel as scale changes. This increases the computational complexity exponentially. However, there exists an algorithm that reduces the computational complexity by a factor of  $2^l$ . The basic idea behind this algorithm is to keep the size of the filter kernels unchanged throughout multiscale analysis. Instead, we partition the input image into four sub-regions so that each region contains only those pixels affected by the filter kernel (without expansion).

Figure 9 illustrates differences between the two approaches: a linear convolution with zero-paddings in the kernel, and an alternative method by splitting the convolved image into four sub-regions for further convolution at finer levels. The former is shown in Figure 9(a). At level 2, zeroes are inserted between the non-zero filter coefficients. Figure 9(b) graphically illustrates the new method. First, the input image is convolved with the original filter kernel at level 1. The convolved image is then divided into four sub-regions such that each region consists of only those pixels processed by the convolution at the next level ( $l = 2$ ). After this splitting, convolutions are then applied to each of the sub-regions independently. To visually describe this, we show within the input image matrix four distinct pixel shapes: dot, triangle, square, and  $\times$ , in such a way that pixels that belong to the same group are identified.

For example, let us compute the value located at coordinate (2,2) in the input image shown in Figure 9(a) marked by a dot. All pixels within the rectangular area from (0,0) to (4,4) will participate in the computation. However, with respect to traditional convolution the zeros in the filter kernel at level 2 effectively mask out all other pixels having shapes other than a dot. Thus pixels that actually take part in the computation are those having a dot shape located at coordinates (0,0), (0,2), (0,4), (2,0), (2,2), (2,4), (4,0), (4,2) and (4,4). The region labeled *I* shown in Figure 9(b) at the level 2 contains exactly those pixels. The convolution is carried out separately in this region and gives us the same result numerically



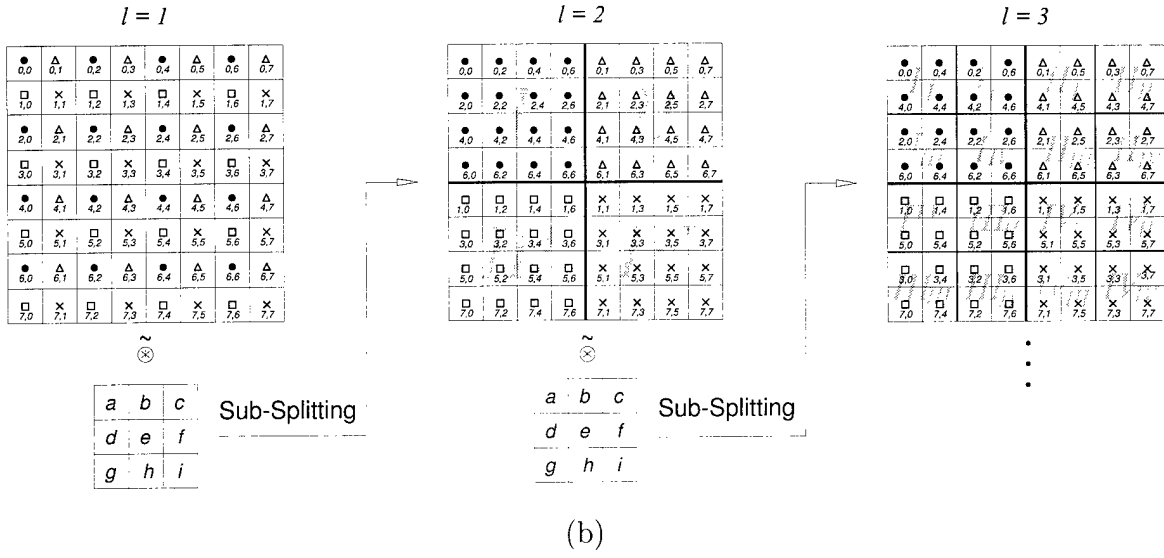
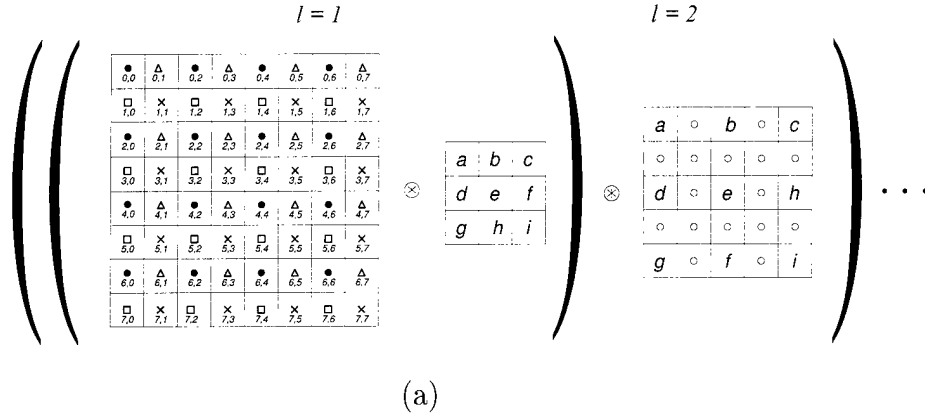


Figure 9: (a) Traditional linear convolution (denoted by  $\otimes$ ) at scale  $l = 1$ ,  $l = 2$ . One zero is padded between non-zero filter coefficients in the filter kernel at scale 2. (b) Alternative approach that avoids zero-paddings between the non-zero filter coefficients. We first convolve the input image with the original filter kernel at the level 1. The convolved image is then divided into four sub-regions, marked in the background as *I*, *II*, *III*, and *IV* with half-grayed tone, each of which holds pixels that are precisely affected by the filter kernel coefficients at the next level  $l = 2$ . Four distinct pixel shapes inside the input image are shown:  $\bullet$ ,  $\triangle$ ,  $\square$  and  $\times$ . Only pixels that have the same shape are processed by convolution at the next level (denoted by  $\tilde{\otimes}$ ).

as the traditional method.

Synthesis is simply the reverse of analysis. In other words, we first merge four sub-regions at a coarser level into a larger region, then convolve it with a filter kernel. Figure 10 shows the synthesis process.

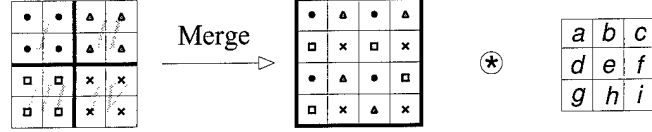


Figure 10: Multiresolution block convolution for image synthesis. Four sub-regions in the input image at a coarser level, marked as *I*, *II*, *III*, and *IV*, are merged into a larger region at a finer level. Then, convolution is applied to the merged region.

For quantitative comparison, let us compute the computational complexity of each method of implementation. For simplicity, we focus on the analysis only. Let the size of the input image and the size of the original filter kernel be  $n$  and  $m$ , respectively. And let us denote the level of analysis by  $l$ . Then, in the traditional linear convolution case, the computational complexity  $C_A$  is

$$C_A = n^2 * (m + 2^l - 2)^2,$$

and in our approach, the computational complexity  $C_B$  is

$$C_B = n^2 * 2^{l-1} * m^2.$$

Normally, the size of the filter kernel is small, so we can assume  $m$  is a constant. The speedup by the new approach is then

$$O\left(\frac{C_A}{C_B}\right) \approx O\left(\frac{2^{2l}}{2^l}\right) = O(2^l).$$

**Algorithm 1** computes the two-dimensional discrete wavelet transform by this more efficient method. The variables,  $n$  and  $m$ , denote the size of the input image and the size of the filter kernel, respectively.  $[FilterKernel]_{X,Y}$  is created by the tensor product of two 1D filter kernels:  $X, Y$ . Thus,

$$[FilterKernel]_{X,Y} = \mathbf{X}' \mathbf{Y}.$$

$[X]_{r,c}^{w,h}$  defines a ROI of size  $w \times h$  starting at  $[r \times w, c \times h]$ . There are three global variables in the algorithm: *LEVEL*, *WIDTH*, *HEIGHT*. By *LEVEL*, we denote the number

---

**Algorithm 1** Two-dimensional discrete dyadic wavelet transform using block splitting method

---

**Require:**  $n$  : dyadic number ( $n > m$ )

**for**  $l = 0$  to  $l < LEVEL$  **do**

$w = \frac{WIDTH}{2^l}$  {  $w$  : width of ROI }

$h = \frac{HEIGHT}{2^l}$  {  $h$  : height of ROI }

**for**  $r = 0$  to  $r < 2^l$  **do**

**for**  $c = 0$  to  $c < 2^l$  **do**

$[ROI]_{r,c}^{w,h} \Leftarrow$  Obtain a region of interest from the input image with size of  $w \times h$  starting at  $[r \times w, c \times h]$

$[W_{2^l}^1 f]_{r,c}^{w,h} \Leftarrow \text{Split} \left( [ROI]_{r,c}^{w,h} * [FilterKernel]_{L,G} \right)$

$[W_{2^l}^2 f]_{r,c}^{w,h} \Leftarrow \text{Split} \left( [ROI]_{r,c}^{w,h} * [FilterKernel]_{G,L} \right)$

$[ROI]_{r,c}^{w,h} \Leftarrow \text{Split} \left( [ROI]_{r,c}^{w,h} * [FilterKernel]_{H,H} \right)$   
{ The process of block splitting is shown in Figure 9(b). }

**end for**

**end for**

**end for**

---



---

**Algorithm 2** Two-dimensional inverse discrete dyadic wavelet transform by merging sub-regions

---

**Require:**  $n$  : dyadic number ( $n > m$ )

**for**  $l = LEVEL - 1$  downto  $l = 0$  **do**

$w = \frac{WIDTH}{2^l}$  {  $w$  : width of ROI }

$h = \frac{HEIGHT}{2^l}$  {  $h$  : height of ROI }

**for**  $r = 0$  to  $r < 2^l$  **do**

**for**  $c = 0$  to  $c < 2^l$  **do**

$[ROI]_{r,c}^{w,h} \Leftarrow$  Merge sub-regions of  $[ROI]_{r,c}^{w,h} * [FilterKernel]_{H,H}^*$

$[W_{2^l}^1 f]_{r,c}^{w,h} \Leftarrow$  Merge sub-regions of  $[W_{2^l}^1 f]_{r,c}^{w,h} * [FilterKernel]_{L,G}^*$

$[W_{2^l}^2 f]_{r,c}^{w,h} \Leftarrow$  Merge sub-regions of  $[W_{2^l}^2 f]_{r,c}^{w,h} * [FilterKernel]_{G,L}^*$

$[ROI]_{r,c}^{w,h} \Leftarrow [ROI]_{r,c}^{w,h} + [W_{2^l}^1 f]_{r,c}^{w,h} + [W_{2^l}^2 f]_{r,c}^{w,h}$   
{ The process of merging of sub-regions is shown in Figure 10. }

**end for**

**end for**

**end for**

---

of levels for the transform, and by *WIDTH* and *HEIGHT* the width and height of the input image, respectively.

**Algorithm 2** describes our algorithm for computing two-dimensional inverse discrete dyadic wavelet transform using the method shown in Figure 10.  $[FilterKernel]_{X,Y}^*$  is the complex conjugate of  $[FilterKernel]_{X,Y}$ .

We show visually in Figure 11 the process of a two dimensional forward wavelet transform using the method described in **Algorithm 1**. In Figure 11(b), we display a series of DC components at levels  $l = 1$ ,  $l = 2$  and  $l = 3$ . Finally, Figure 11(c) shows an enhanced version of the selected ROI shown in Figure 11(a) accomplished by execution on the SPARCstation.

We show in Table 4 the computational costs in time (second) for enhancement between methods of convolution by zero-padding between non-zero filter coefficients, forward and inverse FFT, and our new multiscale approach of splitting and merging sub-regions. Our sample input ROI was  $256 \times 256$  and was computed up to 5 levels of analysis. The size of the original kernels used for the decomposition was  $3 \times 3$  for both horizontal and vertical directional filters. In the reconstruction case, the kernels were  $3 \times 5$  in the horizontal direction and  $5 \times 3$  in the vertical direction.

<i>level</i>	<i>Multiscale Block</i>	<i>Traditional convolution</i>	<i>FFT approach</i>
1	1.18	0.83	13.75
2	1.81	2.87	27.46
3	3.71	7.87	41.05
4	6.29	26.01	54.65
5	13.03	93.93	68.62

Table 4: Computational costs in time (seconds) of splitting (in decomposition) and merging (in reconstruction) schemes versus traditional linear convolution with zero-padding between each non-zero filter coefficients and FFT implementations.

The costs shown in this table were measured on a SPARCstation 20 Model 71 from Sun Microsystems having a 75MHz SuperSPARC-II processor with 1MB SuperCache, 96MB RAM, and 4MB (60ns) SX frame buffer. This table clearly shows that our approach is more efficient than existing convolution schemes and exceeded an FFT by a factor of 5 for five levels of processing.

We have developed a user interface, called **XEnh**, which provides an environment for our algorithm using XIL foundation-level imaging libraries. This program allows users to select a ROI in the main window and carry out a multiscale contrast enhancement algorithm (described in Chapter 2.1) within this ROI. Users are able to set parameters

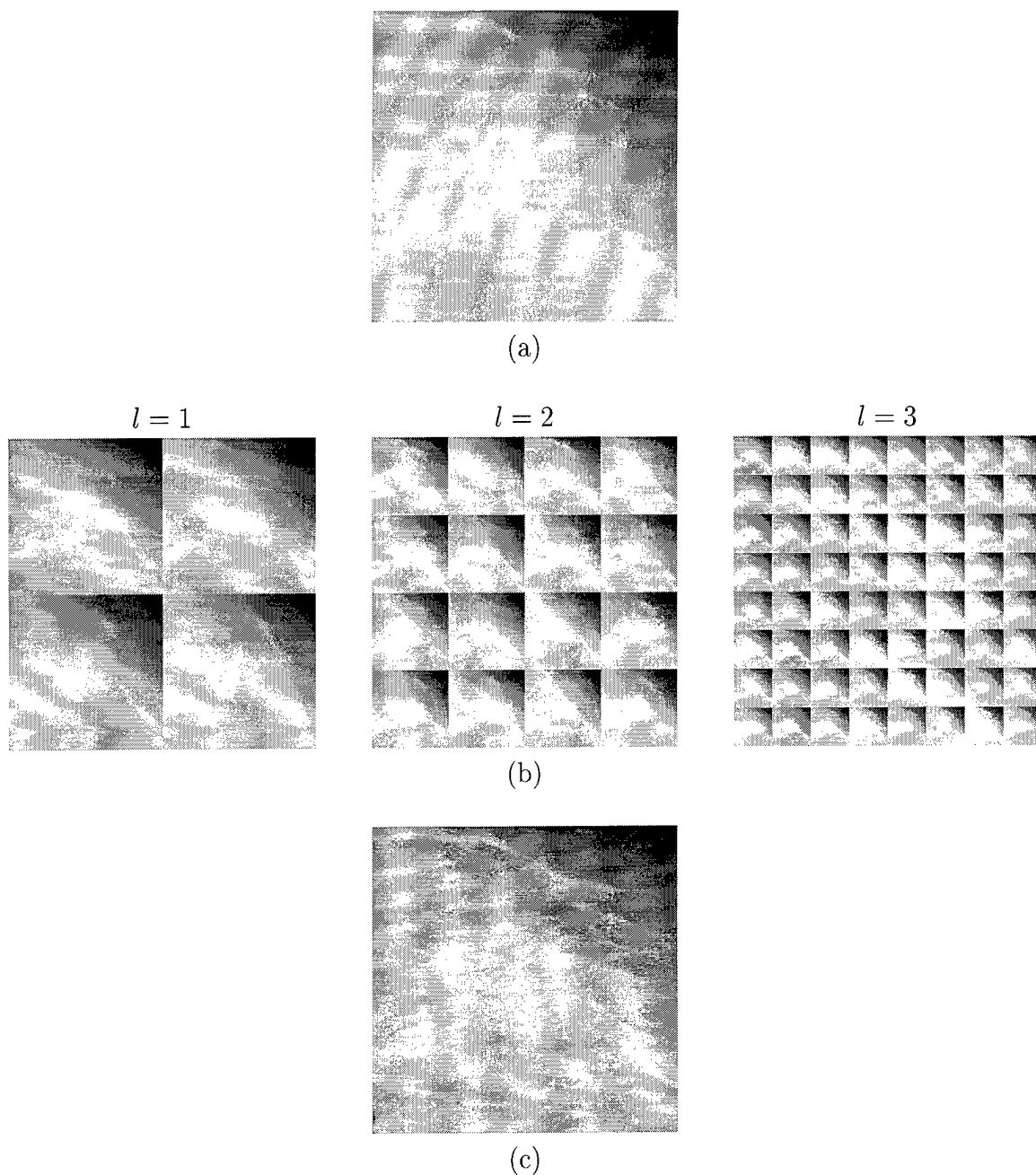


Figure 11: (a) Selected ROI containing an ill-defined mass. (b) DC components at levels  $l = 1$ ,  $l = 2$  and  $l = 3$  after applying **Algorithm 1**. (c) Processed ROI with suspicious area enhanced.

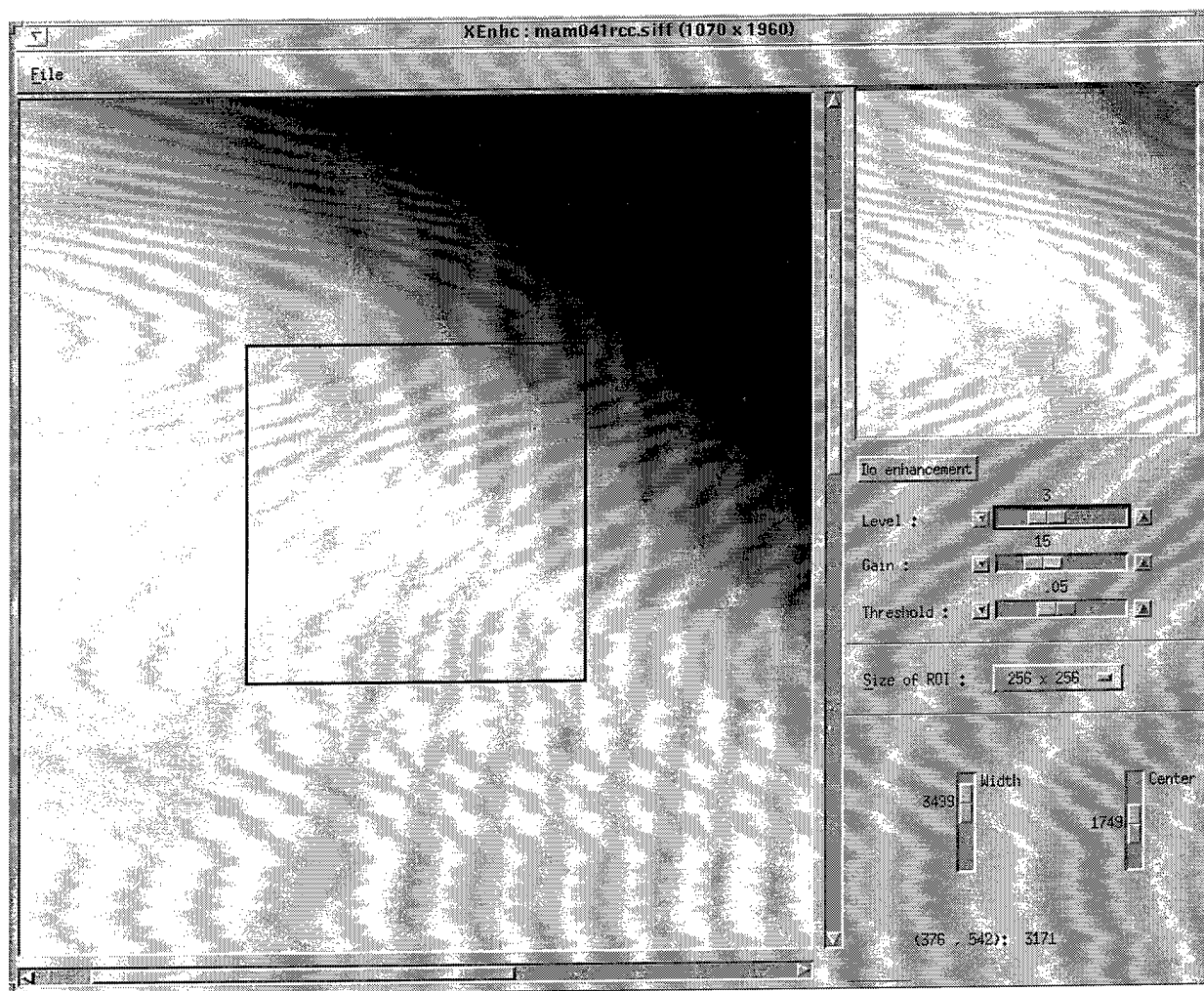


Figure 12: A user interface, **XEnh**, that implements our high-speed parallel algorithm for interactive enhancement.

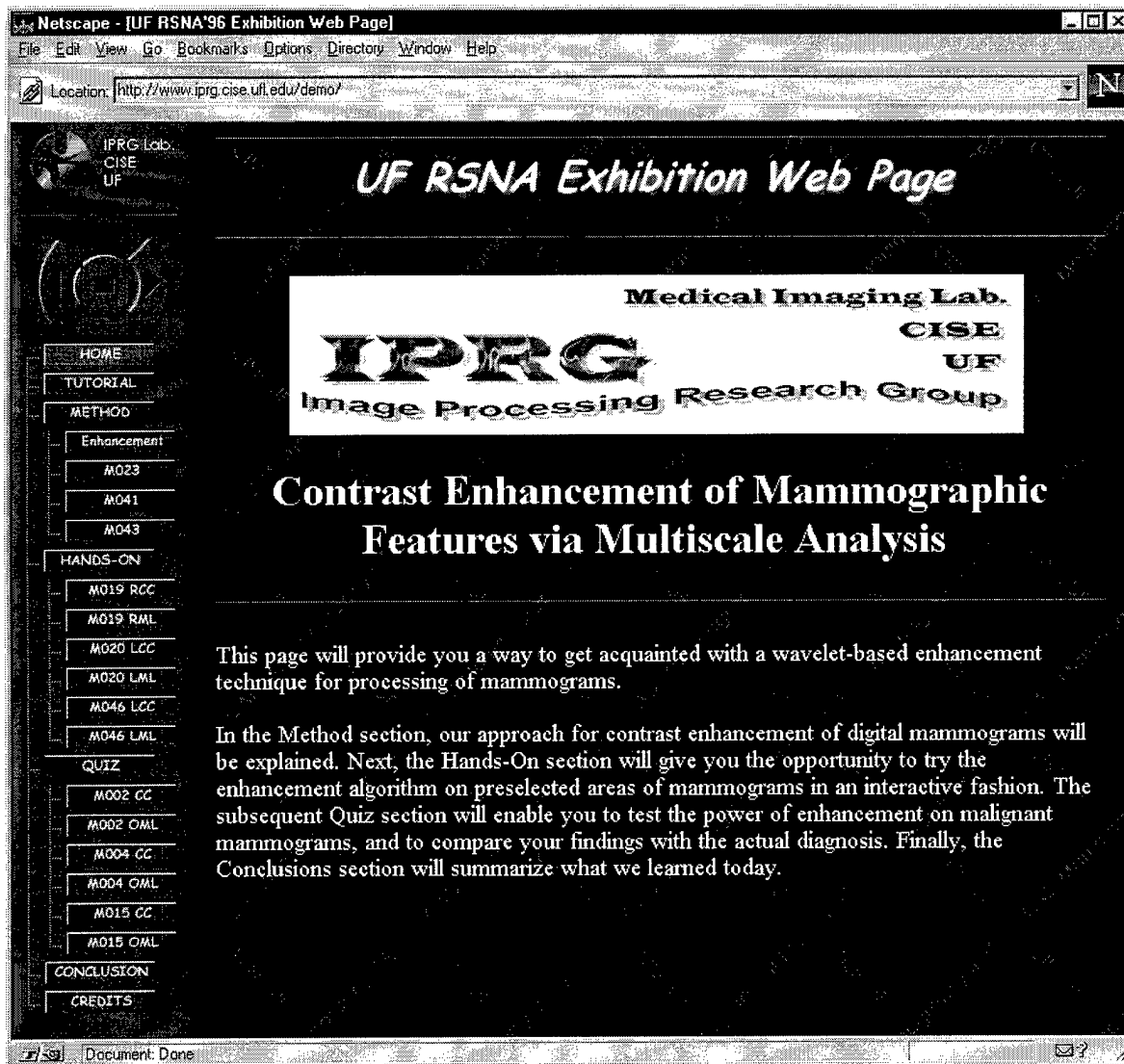
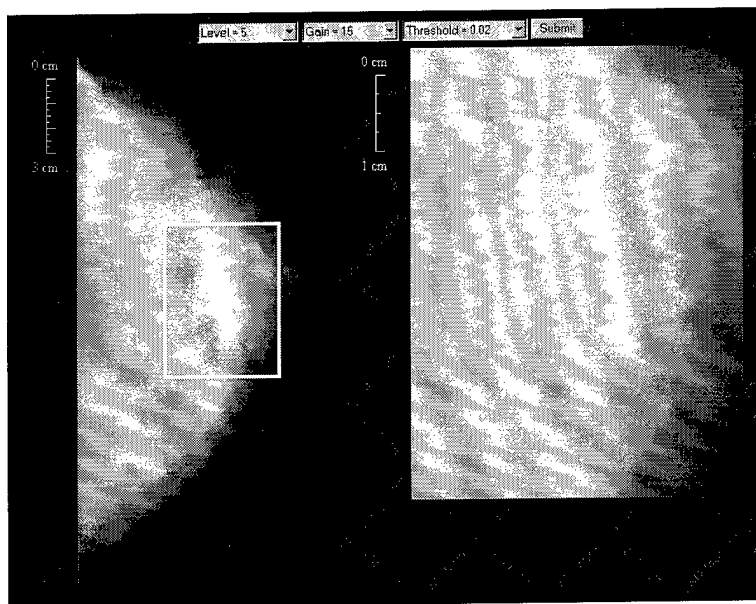
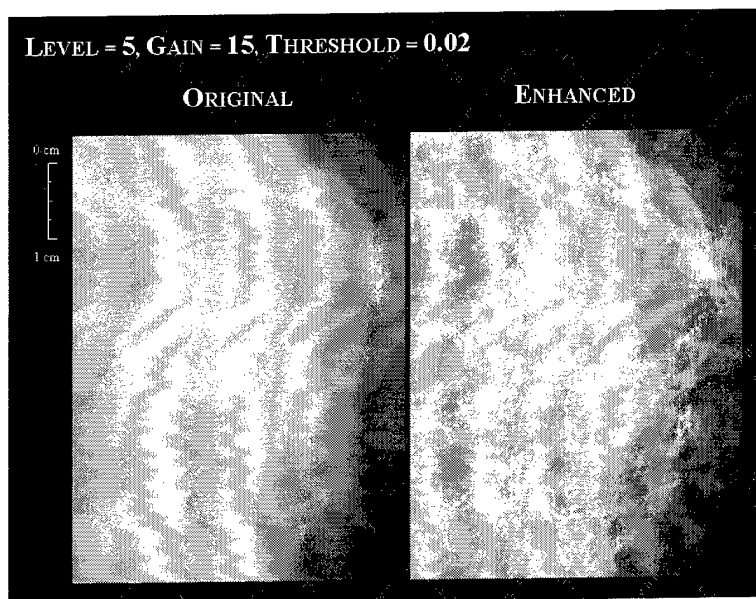


Figure 13: RSNA '96 Exhibition Home-page (<http://www.iprg.cise.ufl.edu/demo>).



(a)



(b)

Figure 14: Hands-on section of the RSNA '96 Exhibition Home-page shown in Figure 13. (a) M019 case with biopsy proven masses. A ROI is preselected. (b) Result of enhancement on the ROI selected in (a).



interactively, such as the level of transform, and values of gain and threshold. The GUI for **XEnh** is shown in Figure 12. The window in the upper right corner of the application is the area where a selected ROI and its enhanced result are displayed.

A web-site (<http://www.iprg.cise.ufl.edu/demo>) shown in Figure 13 has been established, and was originally created for the RSNA (Radiological Society of North America) 1996 exhibition in Chicago, Illinois. In this page, we provide a wavelet-based enhancement technique for processing of mammograms. Sections on this page include “Hands-on” and “Quiz”, where users can process preselected or interactively selected ROIs of mammograms and obtain enhanced results based on the algorithm described in Chapter 2.1.

Figure 14 shows the interface during the process of enhancement for a sample case: M019 RCC, where biopsy proven masses exist. Figure 14(a) shows the case with a preselected ROI. Figure 14(b) displays the result of enhancement on the preselected ROI shown in Figure 14(a).

### **2.2.5 Summary**

We have described a parallel algorithm to accomplish high-speed interactive multiscale processing for enhancement within ROIs in digital mammograms. The implementation relied upon foundation-level libraries, XIL, on top of a SPARCstation’s SX frame buffer. We showed the amount of speedup for the method compared to traditional linear convolution and a conventional FFT approach. Our multiscale approach employing splitting in decomposition and merging in reconstruction was shown to be efficient.

## **2.3 A Continuous Scale Discrete Wavelet Transform for the Detection of Spicular Masses in Mammograms**

### **2.3.1 Introduction**

The Continuous Wavelet Transform (CWT) has been shown to be an effective tool for the analysis of non-stationary signals [8] [41]. A CWT performs a time-scale analysis, zooms in on singularities, captures slow variations and provides a scale-invariant interpretation of shapes.

The importance of scale has been acknowledged in computer vision for a long time. For example, in 1980, Marr and Hildreth developed a multiscale edge detection algorithm modeled on the human visual system [42]. More recently, it has been shown that wavelet theory can provide a solid foundation for multiscale analysis. For example, tracking wavelet transform maxima across scales can allow us to discriminate intensity profiles of

distinct types of edges.

In image processing, Discrete Wavelet Transforms (DWT) are more popular than CWTs because most signals are digitally sampled. Mallat first developed a fast DWT algorithm based on multi-resolution analysis [41]. It computed the wavelet coefficients on a sampled grid  $(a, b)$  defined by  $a = 2^i, b = 2^i k$ , where  $i, k \in \mathcal{Z}$ . Its wavelet representation is complete and concise. The “à trous” algorithm calculates wavelet coefficients at a denser sample grid,  $a = 2^i, b = k$  [13]. An important advantage of the “à trous” algorithm for image analysis is that it is shift-invariant.

Visual features occur at distinct scales. For feature analysis, it is desirable to detect features at an arbitrary scale. M. Unser *et al.* [43] [44] have proposed several fast and efficient algorithms for calculating wavelet coefficients at arbitrary scales. In these algorithms,  $Q$  scales were inserted into each octave to obtain a denser sampling of scale. The wavelets at each of the  $Q$  scales in the first octave were approximated by their projections in scale space, and called auxiliary wavelets. Since the auxiliary wavelets lie in scale space, auxiliary wavelets at the next higher octave were obtained by using the standard two-scale relation.

In this chapter, we introduce a special type of mother wavelet that allows for an arbitrary sampling of the scale parameter. We discuss how to create such a wavelet, which is an approximation of a known mother wavelet. We also develop a method to reconstruct an original image from its wavelet coefficients generated from an arbitrary scale analysis. We demonstrate the importance of arbitrary scale analysis by applying the method to problems of detecting breast mass lesions of arbitrary size in digitized mammograms.

The chapter is organized as follows. In Section 2.3.3, we first derive the Continuous Scale Discrete Wavelet Decomposition from the CWT. In Section 2.3.4, we show how to reconstruct a signal from its wavelet coefficients at an arbitrary scale  $a$  and at scales  $2^{-i}a$ , where  $i \in \mathcal{Z}$ ,  $2^{-i}a > a_0$  and  $a_0$  is some initial scale of analysis. In Section 2.3.5, we discuss an algorithm using wavelets that are  $d$ th derivatives of B-spline wavelets [7] [31]. We then define a general form of filters to implement our multiscale analysis and synthesis in the frequency domain. In Section 2.3.6, we show the results of applying our algorithm to difficult mammography cases of breast cancer. The ability to find and process features at an arbitrary scale is shown to have significant advantages over standard dyadic algorithms. In one case, the algorithm detected the center of a mass that could not be seen using conventional window and leveling techniques with 12-bit contrast resolution. In Section 2.3.7, we discuss the relation of our decomposition scheme with the oblique projection scheme proposed by M. Vrhel, M. Unser *et al.* [43] and show that they are identical for a limiting case. Finally, in Section 2.3.8, we summarize our results.

### 2.3.2 Definitions and Notation

In this chapter, a letter in lower case represents a function or a sequence in the time or spatial domain. Its upper case represents the corresponding frequency response, i.e.  $p(k) \leftrightarrow P(\omega)$ ,  $\psi(t) \leftrightarrow \Psi(\omega)$ .  $\psi_a(t) = \psi(\frac{t}{a}) \leftrightarrow \Psi_a(\omega) = \Psi(a\omega)$ . Given a sequence  $x(n)$ , a function  $x^c(t)$  is reconstructed from  $x(n)$  using an ideal lowpass filter.

The symbol “\*” will be used for three distinct types of convolutions: linear convolution, mixed convolution and discrete convolution. For functions  $f$  and  $g$  defined on  $\mathcal{R}$ , “\*” denotes traditional convolution

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau.$$

The mixed convolution between a sequence  $b(k)_{k \in \mathcal{Z}}$  and a function  $f$  defined on  $\mathcal{R}$  is

$$(b * f)(t) = \sum_{k=-\infty}^{\infty} b(k)f(t - k).$$

The discrete convolution between two sequences  $a$  and  $b$  is defined as

$$(a * b)(n) = \sum_{k=-\infty}^{\infty} a(k)b(n - k).$$

Whenever it exists, the convolution inverse  $b^{-1}$  of a sequence  $b$  is defined to be

$$(b^{-1} * b) = \delta(k).$$

Function  $\beta^n(t)$  is Schoenberg’s central B-spline of degree  $n$ , which is obtained from the  $(n + 1)$ -fold convolution of a unit rectangular pulse:

$$\beta^n(t) = \sum_{j=0}^{n+1} \frac{(-1)^j}{n!} \binom{n+1}{j} \left[ t + \frac{n+1}{2} - j \right]_+^n, \quad t \in \mathcal{R},$$

where  $[t]_+ = \max\{0, t\}$ . The Fourier transform of  $\beta^n(t)$  is given by  $\hat{\beta}^n(\omega) = \left( \frac{\sin(\frac{\omega}{2})}{\frac{\omega}{2}} \right)^{n+1}$ . Discrete B-spline sequences are obtained by sampling the continuous B-spline functions,  $b^n(k) = \beta^n(t)|_{t=k}$  for  $k \in \mathcal{Z}$ . Note  $(b^n)^{-1}$  is stable for any value of  $n$  [45].

Finally, the function  $Rect(\omega)$  is defined as

$$Rect(\omega) = \begin{cases} 1 & \text{when } |\omega| \leq \pi \\ 0 & \text{when } |\omega| > \pi. \end{cases}$$

### 2.3.3 A Continuous Scale Discrete Wavelet Transform

A wavelet transform is a linear operation that projects a signal onto a set of basis functions, which are dilated versions of some mother wavelet. The mother wavelet is a function  $\psi \in L^2$  that satisfies  $\int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < +\infty$ . This requires that the wavelet has sufficient decay, and that  $\int_{-\infty}^{\infty} \psi(t) dt = 0$ . The continuous wavelet transform of a function at scale  $a$  and shifting parameter  $b$  is defined as

$$\mathbf{W}f(a, b) = \int_{-\infty}^{\infty} f(t) \psi \left( \frac{b-t}{a} \right) dt.$$

Of course, it has been shown that a wavelet transform satisfies energy conservation and can be perfectly reconstructed from its wavelet transform coefficients [8] [46]. The Continuous Scale Discrete Wavelet Transform (CSWDT) is defined as the CWT sampled along the shifting parameter  $b$ ,

$$\mathbf{W}f(a, n) = \int_{-\infty}^{\infty} f(t) \psi \left( \frac{n-t}{a} \right) dt. \quad (60)$$

In practice, data is acquired digitally. Furthermore, we assume that a signal  $f(t)$  is sampled at its Nyquist rate  $f_s$  so that it may be recovered from its discrete samples. Since  $f(t)$  is band-limited,  $\mathbf{W}f(a, b)$  is also band-limited, and can be recovered from its samples  $\mathbf{W}f(a, n)$ .

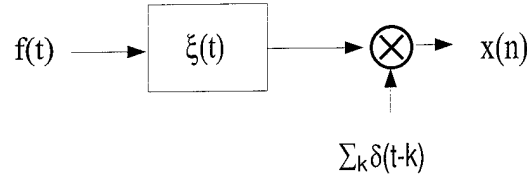


Figure 15: Data acquisition processing.

Processing for data acquisition is shown in Figure 15 ( $\xi(t)$  is an anti-aliasing filter). Since the sampled signal  $x(n)$  is of finite resolution, this puts a limit on the finest scale that can be computed. Let us introduce a compactly supported scaling function  $\phi(t)$  that satisfies the following conditions [43] :

1. Riesz basis condition:  $A \leq \sum_{k \in \mathbb{Z}} |\Phi(\omega + 2k\pi)|^2 \leq B$ , where  $A$  and  $B$  are strictly positive constants;
2. Order property:  $\Phi(0) = 1, \Phi^{(m)}(2k\pi) = 0, k \in \mathbb{Z}, k \neq 0$ , for  $m = 0, \dots, N-1$ ;
3. Two-scale relation:  $\phi\left(\frac{t}{2}\right) = \sum_{k \in \mathbb{Z}} h(k) \phi(t-k)$ ;

4. Frequency property:  $\Phi(\omega) \neq 0$ , for  $|\omega| \leq \pi$ , and  $\Phi(\omega)$  should be very small for  $|\omega| > \pi$ .

Constraint 1 implies that  $\phi(t - k)$  is a Riesz basis of the subspace

$$V_\phi = \left\{ f(t) = \sum_{k \in \mathbb{Z}} c(k) \phi(t - k) \quad c \in \ell_2 \right\} \quad \text{for } f(t) \in V_\phi$$

and that  $V_\phi$  is a well defined (closed) subspace of  $\mathcal{L}_2$  [47]. Property 2 implies that  $\phi(t)$  reproduces all polynomials of degree  $N-1$  [48]. Constraint 3 is the well-known two-scale relation. Condition 4 shows that  $\phi(t)$  can be an anti-aliasing filter.

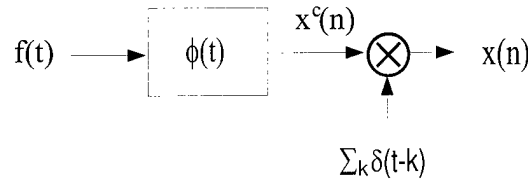


Figure 16: Initial conditions.

Typically, the anti-aliasing filter in the acquisition device is unknown. For simplicity, we assume that

$$x(n) = \int_{-\infty}^{\infty} f(t) \phi(n - t) dt, \quad (61)$$

which is shown in Figure 16. This is the initial condition typically found in the traditional DWT analysis literature. When the impulse response of the anti-aliasing filter (i.e. ideal low-pass filter) is known, a pre-filter and a post-filter are used as shown in Figure 17 (Please See Appendix A.1 for details).

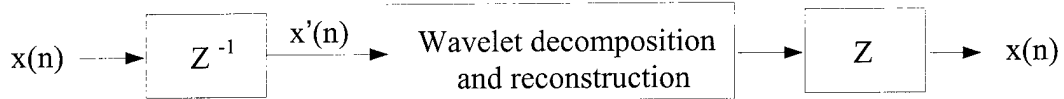


Figure 17: Wavelet processing with alternative initial conditions.

The smoothing operator  $S_a$  is defined as  $S_a f(n) = \int_{-\infty}^{\infty} f(t) \phi\left(\frac{n-t}{a}\right) dt$ . Then  $S_1 f(n) = x(n)$ , where  $a = 1$  is the finest scale at analysis. As  $a$  increases,  $S_a f(n)$  becomes a coarser representation of  $f(t)$ . When  $a$  goes to infinity,  $S_a f(n)$  approaches a constant equal to the average value of  $f(t)$ . In practice, the length of  $x(n)$  defines the coarsest scale.

**Lemma 2.1** Given  $x(n)$ , there is one and only one  $f(t)$  such that

$$x^c(t) = f * \phi(t). \quad (62)$$

**Proof:** Since  $X^c(\omega)$  is band-limited to  $[-\pi, \pi]$  and  $\Phi(\omega) \neq 0$  for  $\omega \in [-\pi, \pi]$ ,  
 $F(\omega) = \frac{X^c(\omega)}{\Phi(\omega)}$ . ■

**Theorem 1** Given  $x(n) = x^c(t)|_{t=n}$ , if

1.  $\tilde{\Psi}_a(\omega)$  is band-limited between  $[-\pi, \pi]$  when  $a > s_0$ ;
2.  $a > s_0$ .

Then, the CSDWT can be calculated by

$$\mathbf{W}_a f(a, n) = \sum_{k \in \mathcal{Z}} p(a, k) x(n - k), \quad (63)$$

where

$$P^c(a, \omega) = \frac{\tilde{\Psi}_a(\omega)}{\Phi(\omega)} \quad (64)$$

and

$$P^c(a, \omega) = \int_{-\infty}^{\infty} p^c(a, t) e^{-j\omega f_s t} dt.$$

**Proof:** Note that  $\tilde{\psi}_a(t)$  and  $p^c(a, t)$  are band-limited at  $[-\pi, \pi]$ . And  $p(a, n) = p^c(a, t)|_{t=n}$ .  
 From Equation (64),

$$\tilde{\Psi}_a(\omega) = P^c(a, \omega) \Phi(\omega) \quad (65)$$

$$\tilde{\psi}_a(t) = \int_{-\infty}^{\infty} p^c(a, u) \phi(t - u) du. \quad (66)$$

If we substitute Equation (66) into the previous CSDWT definition (Equation (60)), we obtain

$$\begin{aligned} \mathbf{W}_a f(a, n) &= \int_{-\infty}^{\infty} f(t) \tilde{\psi}_a(n - t) dt \\ &= \int_{-\infty}^{\infty} f(t) \int_{-\infty}^{\infty} p^c(a, u) \phi(n - t - u) du dt \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(t) \phi(n - t - u) dt \right) p^c(a, u) du \\ &= \int_{-\infty}^{\infty} x^c(n - u) p^c(a, u) du. \end{aligned} \quad (67)$$

Since  $x^c(t)$  and  $p^c(a, t)$  are band-limited to  $[-\pi, \pi]$ , we need to show that

$$\int_{-\infty}^{\infty} x^c(n-u)p^c(a, u)du = \sum_{k \in \mathbb{Z}} p(a, k)x(n-k). \quad (68)$$

If we look in the frequency domain, we see that

$$\mathcal{F}\left\{\int_{-\infty}^{\infty} x^c(n-u)p^c(a, u)du\right\} = (X^c(\omega)P^c(a, \omega)) * 2\pi\delta_{2\pi}(\omega).$$

Since  $x^c(t)$  and  $p^c(a, t)$  are band-limited to  $[-\pi, \pi]$ ,

$$\begin{aligned} \mathcal{F}\left\{\int_{-\infty}^{\infty} x^c(n-u)p^c(a, u)du\right\} &= (X^c(\omega) * 2\pi\delta_{2\pi}(\omega))(P^c(a, \omega) * 2\pi\delta_{2\pi}(\omega)) \\ &= X(\omega)P(a, \omega). \end{aligned}$$

Therefore, we can claim that Equation (68) is indeed true. Substituting Equation (68) into Equation (67), we obtain

$$\mathbf{W}_a f(a, n) = \sum_{k \in \mathbb{Z}} p(a, k)x(n-k).$$

■

Thus,  $p(a, k)$  of length  $N$  can be calculated as follows:

1. For  $\omega \in [-\pi, \pi]$ , Evaluate  $P(a, \omega) = P^c(a, \omega) = \frac{\tilde{\Psi}_a(\omega)}{\Phi(\omega)}$  at  $\omega = n\frac{\pi}{N}$ , where  $-N \leq n < N$ ;
2. Note that  $P(a, \omega)$  is a  $2\pi$  periodic function, calculate  $P(a, \omega)$  for  $\omega \in [0, 2\pi)$ ;
3. Calculate  $p(a, k)$  from the **IFFT** of  $P(a, \omega)$ .

Usually, a wavelet  $\psi(t)$  is not compact in the frequency domain. However,  $\Psi(\omega)$  is local and we can select an  $s_0$  such that  $\tilde{\Psi}_{s_0}(\omega) = \Psi_{s_0}(\omega)Rect(\omega)$  has an acceptable approximation error onto the scaling space. The approximation error is defined as the ratio of the energy of  $\Psi(a\omega)$  outside  $[-\pi, \pi]$  to the total energy. It decreases as  $s_0$  increases.

Next we show that our  $\tilde{\psi}_{s_0}(t)$  is also a mother wavelet. Since  $\psi(t)$  satisfies the “admissibility condition”,

$$\begin{aligned} \int_{-\infty}^{\infty} |\omega|^{-1} |\Psi(\omega)|^2 d\omega &< \infty \\ \int_{-\infty}^{\infty} |s_0\omega|^{-1} |\Psi(s_0\omega)|^2 d(s_0\omega) &< \infty \\ \int_{-\infty}^{\infty} |\omega|^{-1} |\Psi(s_0\omega)|^2 d\omega &< \infty \\ \int_{-\pi}^{\pi} |\omega|^{-1} |\Psi(s_0\omega)|^2 d\omega &< \infty \\ \int_{-\infty}^{\infty} |\omega|^{-1} |\tilde{\Psi}(s_0\omega)|^2 d\omega &< \infty. \end{aligned} \quad (69)$$

Therefore,  $\tilde{\psi}_{s_0}(t)$  satisfies the “admissibility condition”. Wavelet coefficients at arbitrary scale  $a = ss_0$  can be calculated directly from Equation (63). However, wavelet coefficients at scale  $a = 2^i s_0$  can be calculated efficiently using the fast algorithm discussed in the next section.

### 2.3.4 An Efficient Algorithm for Analysis and Synthesis

When there is knowledge about the best scale in which to detect a certain feature, it is desirable to identify the feature at that scale, enhance the feature and reconstruct the signal. We now derive a computationally efficient scheme to decompose a signal onto scale ( $a = 2^L s_0$ ) and reconstruct the signal at that scale.

**Theorem 2** Let  $\tilde{\psi}_{s_0}(t) = \sum_{k \in \mathbb{Z}} p(s_0, k) \phi(t - k)$ ,

$$\tilde{\Psi}(2^L s_0 \omega) = P^c(s_0, 2^L \omega) \left( \prod_{k=0}^{L-1} H(2^k \omega) \right) \Phi(\omega).$$

**Proof:** Since

$$\tilde{\Psi}(s_0 \omega) = P(s_0, \omega) \Phi(\omega), \quad (70)$$

we can substitute  $\omega$  with  $2^L \omega$ ,

$$\tilde{\Psi}(2^L s_0 \omega) = P(s_0, 2^L \omega) \Phi(2^L \omega). \quad (71)$$

Then, using the two scale relation to simplify  $P(s_0, 2^L \omega) \Phi(2^L \omega)$ ,

$$\begin{aligned} P(s_0, 2^L \omega) \Phi(2^L \omega) &= P(s_0, 2^L \omega) H(2^{L-1} \omega) \Phi(2^{L-1} \omega) \\ &= P(s_0, 2^L \omega) \left( \prod_{k=0}^{L-1} H(2^k \omega) \right) \Phi(\omega). \end{aligned} \quad (72)$$

Substituting Equation (72) into Equation (71),

$$\tilde{\Psi}(2^L s_0 \omega) = P^c(s_0, 2^L \omega) \left( \prod_{k=0}^{L-1} H(2^k \omega) \right) \Phi(\omega). \quad (73)$$

■

If we let  $K(\omega)$  be a  $2\pi$  periodic function. For  $\omega \in [-\pi, \pi]$ ,

$$K(\omega) = \frac{1 - |H(\omega)|^2}{P(s_0, \omega)}. \quad (74)$$





Its wavelet is written as

$$\Psi(\omega) = (j\omega)^d \left( \frac{\sin(\frac{\omega}{4})}{\frac{\omega}{4}} \right)^{n+d} \quad (78)$$

and

$$\psi(\frac{t}{2}) = \frac{d^d \beta^{n+d-1}(t)}{dt^d} = \sum_{k \in \mathcal{Z}} g(k) \beta^{n-1}(t-k). \quad (79)$$

The two-scale relation becomes

$$\Phi(2\omega) = e^{-jt_0\omega} H(\omega) \Phi(\omega), \quad (80)$$

and

$$H(\omega) = e^{jt_0\omega} \cos^n \left( \frac{\omega}{2} \right), \quad (81)$$

where  $t_0 = \frac{1}{2}$  when  $n$  is odd and level  $L = 0$ . Otherwise,  $t_0 = 0$ .  $\psi(t)$  is a  $d$ -th derivative of a spline of order  $n + d - 1$  and  $\phi(t)$  is a spline function of order  $n - 1$ . The  $e^{-jt_0\omega}$  term is introduced because central B-splines have knots at  $i \in \mathcal{Z}$  for  $n$  even and at  $i + \frac{1}{2}$  for  $n$  odd.

In our fast decomposition and reconstruction algorithm, for  $\omega \in [-\pi, \pi]$ , we can write

$$P(s_0, \omega) = e^{jt_0\omega} \frac{j^d 2^{n+2d} \sin^{n+d}(\frac{s_0\omega}{4})}{s_0^n \sin^n(\frac{\omega}{2})} \quad (82)$$

$$H(\omega) = e^{jt_0\omega} \cos^n \left( \frac{\omega}{2} \right) \quad (83)$$

$$K(\omega) = \frac{1 - |H(\omega)|^2}{P(s_0, \omega)}. \quad (84)$$

Where,  $t_1 = \frac{1}{2}$  when  $d$  is odd and level  $L = 0$ , and  $t_1 = 0$  otherwise. Notice that when  $s_0 = 2$ , it is easy to show that the filters  $P(s_0, \omega)$ ,  $H(\omega)$  and  $K(\omega)$  are the same as the traditional DWT filters  $G(\omega)$ ,  $H(\omega)$  and  $K(\omega)$ .

As an example, we constructed and applied a spline wavelet with  $d = 2$  and  $n = 3$  to decompose a sample of white noise onto scale 14.25. We then reconstructed the signal from the same scale. Multiscale coefficients of the decomposition are shown in Figure 19. Remarkably, the reconstruction error was less than  $10^{-14}$ , which was due principally to the floating point accuracy of our digital computer, a Sun SparcStation 20.

In the next section, we apply this algorithm to a digital radiograph of a phantom and sample digitized mammogram cases.

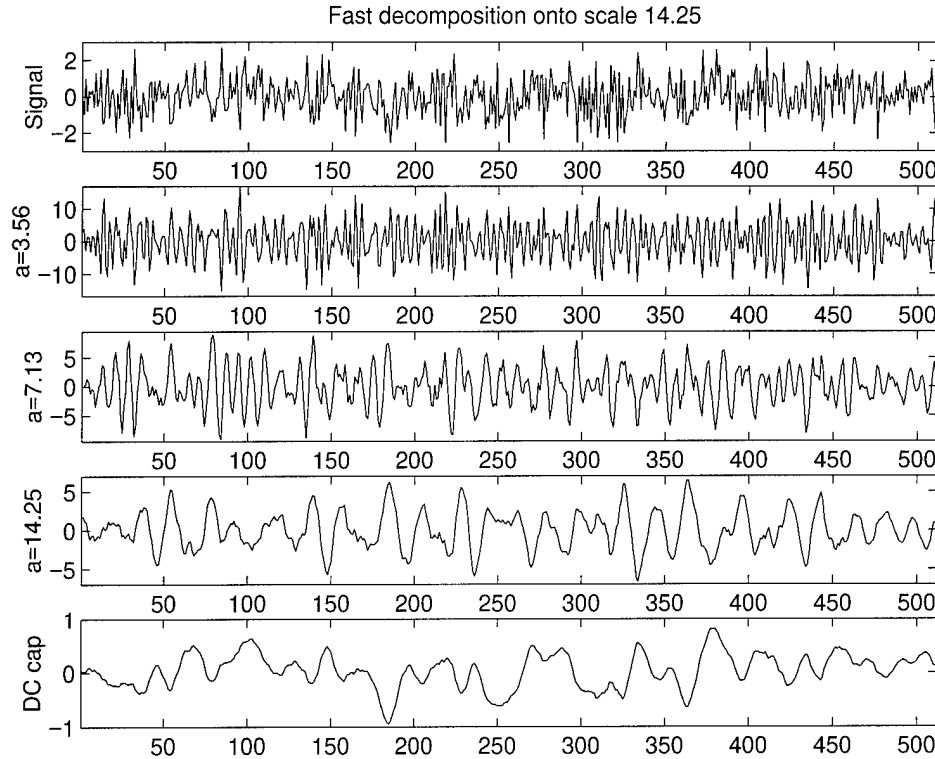


Figure 19: Sample of fast decomposition algorithm of white noise at scale 14.25.

### 2.3.6 Application of the CSDWT Algorithm

X-ray images of an RMI model 156 phantom (Radiation Measurements Inc., Middleton, WI) were investigated. The RMI phantom contains five masses and six fibrils which mimic two objects of interest in mammography. The radiographic image quality of this phantom, as well as the total number of objects which are deemed to be visible, form a key component of the mammography accreditation program administered by the American College of Radiology (ACR). The RMI phantom is designed so that at least one object (i.e. smallest mass) in each category is not visible, and one object is at the borderline of the visibility threshold [52].

Radiographic images of the RMI phantom were obtained using a Digital Spot Mammography (DSM) system attached to a Lorad Breast Biopsy System (Lorad Corporation, Danbury, CT). Due to the limit of the detector area, only regions containing masses at the right-lower corner were captured (with technique factors of 22 kVp and 112 mAs) as shown in Figure 20(a). The smallest mass in the schematic representation of the insert of the phantom (Figure 20(e)) can not be seen using conventional window and leveling at 12-bit contrast resolution.

In this study, the wavelet applied was a second derivative of a spline of order 5 ( $n=4$ ,

$d=2$ ). The scaling function was a cubic spline. The 2-D wavelet filters were derived from the 1-D design using a rotated circularly symmetric window [53]. The minimums of wavelet coefficients computed at scale 53.8 are shown in Figure 20(b). A soft threshold 0.01 was then applied. Note that all the masses were correctly detected. Figures 20(c) and 20(d) show that a dyadic wavelet fails to detect all the masses within the phantom.

Mammograms obtained from our local database containing lesions of varying density are described in Table 5.

Case	Shape	Margins	Density	Size	Biopsy	Difficulty
rml041	Irregular	Indistinct	Equal	12mm	Malignancy	subtle
lcc046	Irregular	Indistinct	High	10mm	Malignancy	Mod. Difficult
lcc002	Irregular	Spiculated	High	10mm	Malignancy	Difficult
lml015	Round	Obscured	High	15mm	Malignancy	Difficult

Table 5: Description of mammograms

Original images rml041 and lcc046 are shown in Figure 21. In each case, the scale space was systematically searched (interactively) to find an optimum scale for detection. Then, chains of wavelet coefficient minima at the selected scale were labeled and classified as shown in Figure 22. For proper labelling, the set of maxima were first thinned to be exactly 1 pixel wide. The labelling process first classified each maxima as one of the following:

1. JUNCTION: If a minimum had three or more (8-connected) neighbors;
2. TERMINAL: If a minimum had zero or one neighbor, or exactly two neighbors at least one of which is a JUNCTION;
3. EDGE: If a minimum had exactly two neighbors, neither a JUNCTION.

A set of contour-level labels were generated, using the TERMINALS to find endpoints. When all TERMINALS had been processed, any EDGES not labelled were linked into contour-level labels. The JUNCTIONs were then processed to produce group-level labels, each representing a set of 8-connected chains. Each contour was examined to determine if it should be split to form chain-level labels. Finally, local standard deviation of the gray level intensity values for all pixels in each contour was then calculated. Each chain was assigned to a Group as shown in Figure 23. Note that chains projecting from the center of the mass, follow the spicular distortion near the mass. In each case, the algorithm detected each mass, at scale 124.8 and scale 112, respectively. However, for rml041, the detection algorithm also picked up an object in the right-lower corner, which looked quite similar to

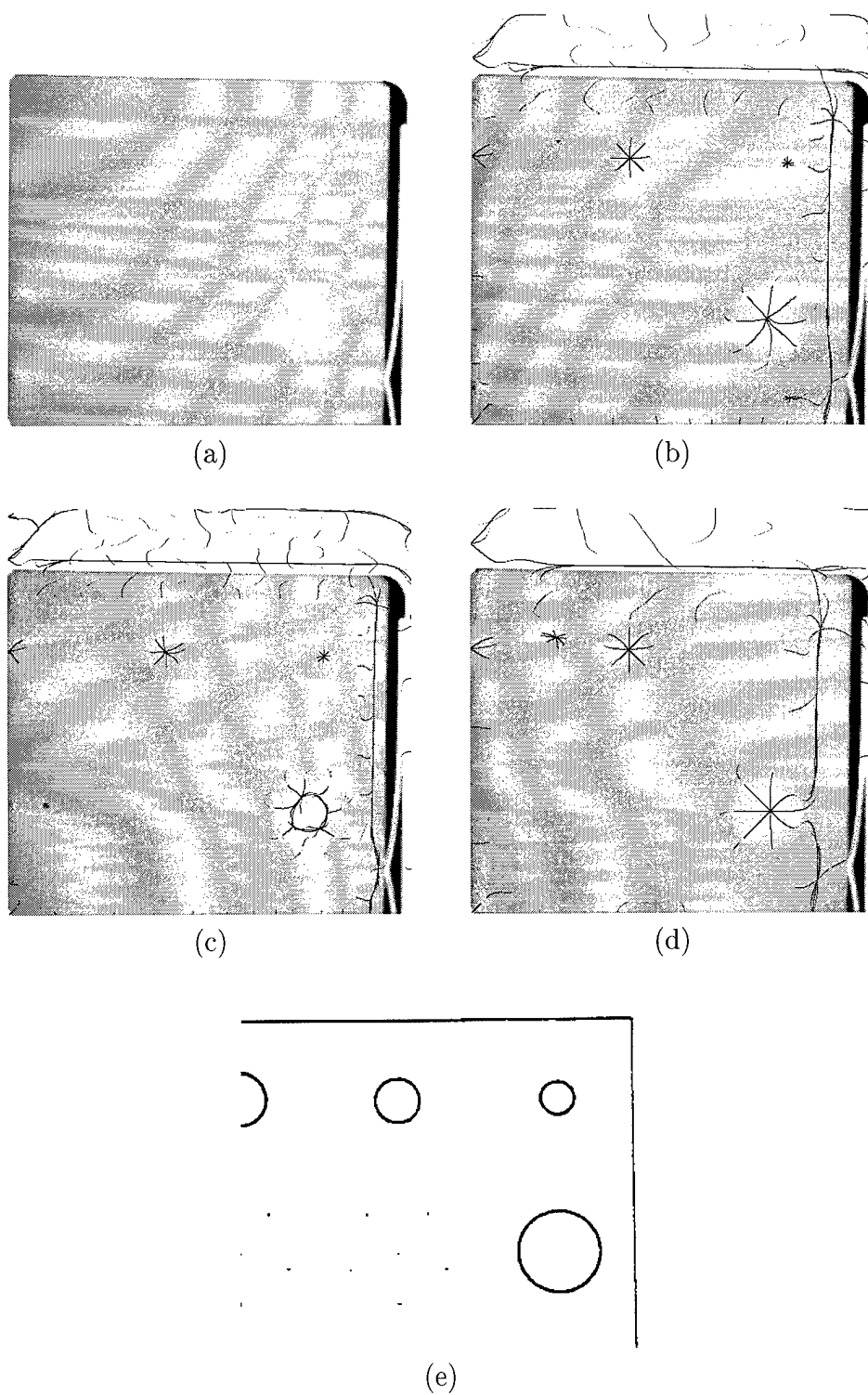
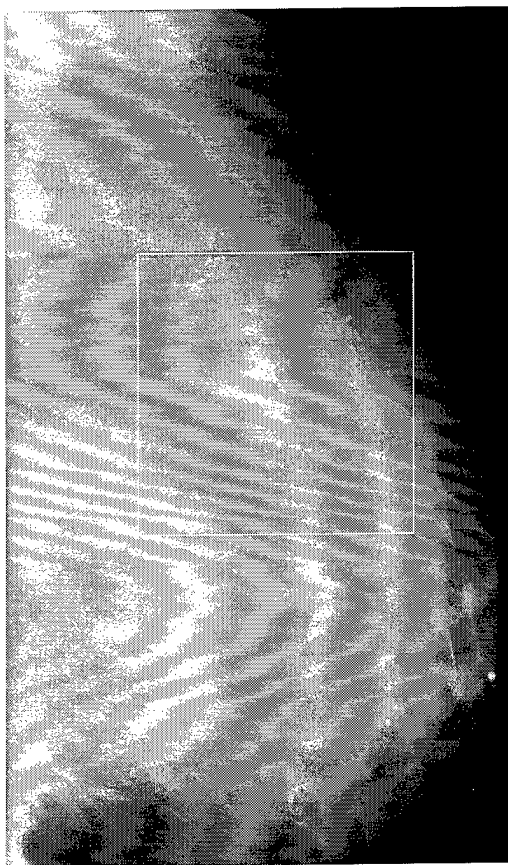
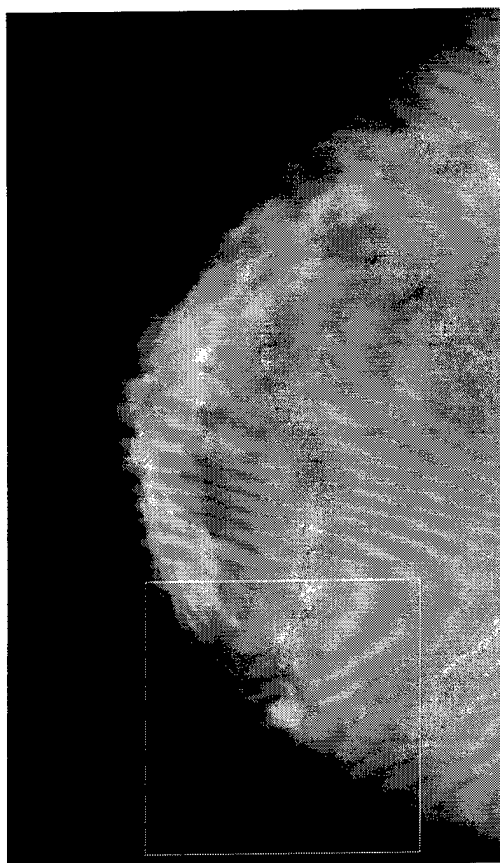


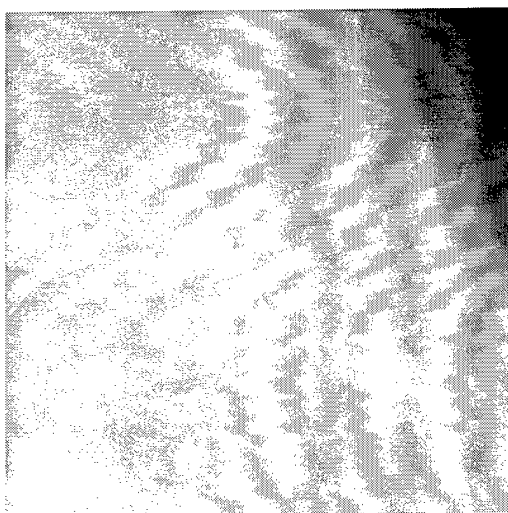
Figure 20: (a) X-ray image of the RMI156 phantom; (b) Coefficient minima obtained at scale 53.8; (c) Coefficient minima obtained at scale 32; (d) Coefficient minima obtained at scale 64; (e) Schematic representation of mammographic features within the phantom.



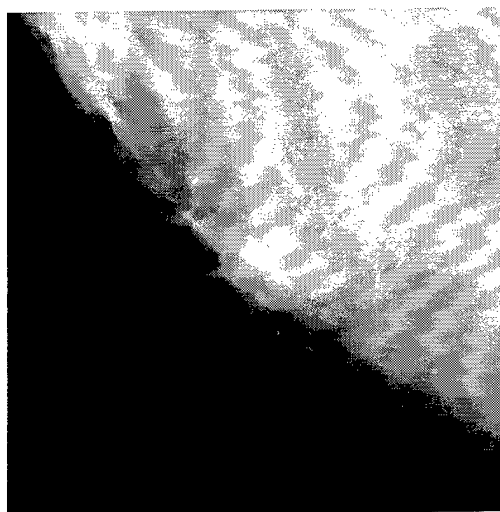
(a)



(b)



(c)



(d)

Figure 21: (a) Mammogram rml041; (b) Mammogram lcc046; (c) ROI cropped from lcc041; (d) ROI cropped from lcc046.

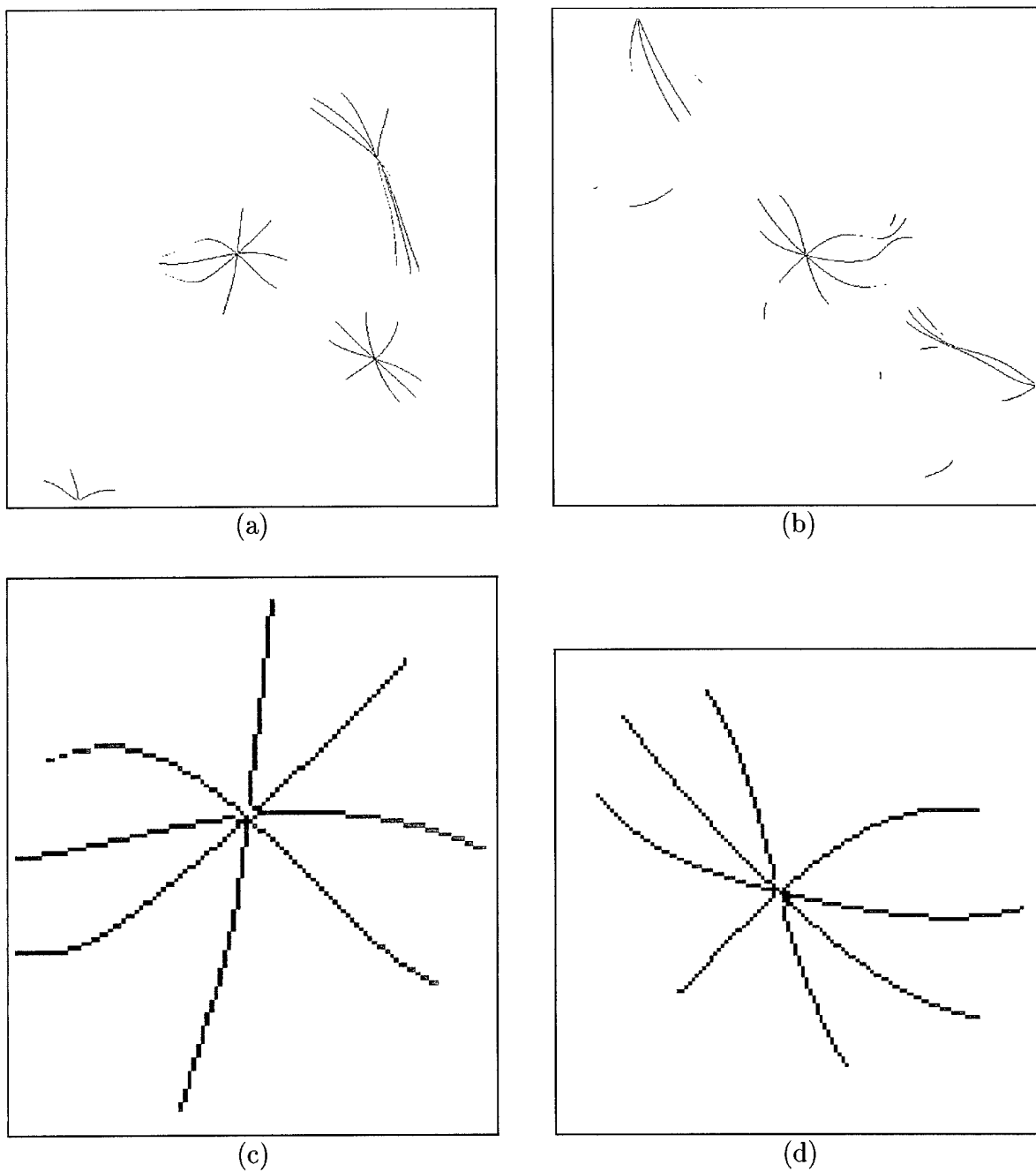


Figure 22: (a) Labeling of wavelet coefficient minima within selected ROI from rml041 (soft-thresholding applied). "JUNCTIONs" are shown in BLUE, "TERMINALs" in GREEN, and "CONTOURs" in red; (b) Labeling of wavelet coefficient minima within ROI from lcc046; (c) 4X magnification of the center cluster representing a mass shown in (a); (d) 4X magnification of the center cluster representing a mass shown in (b).

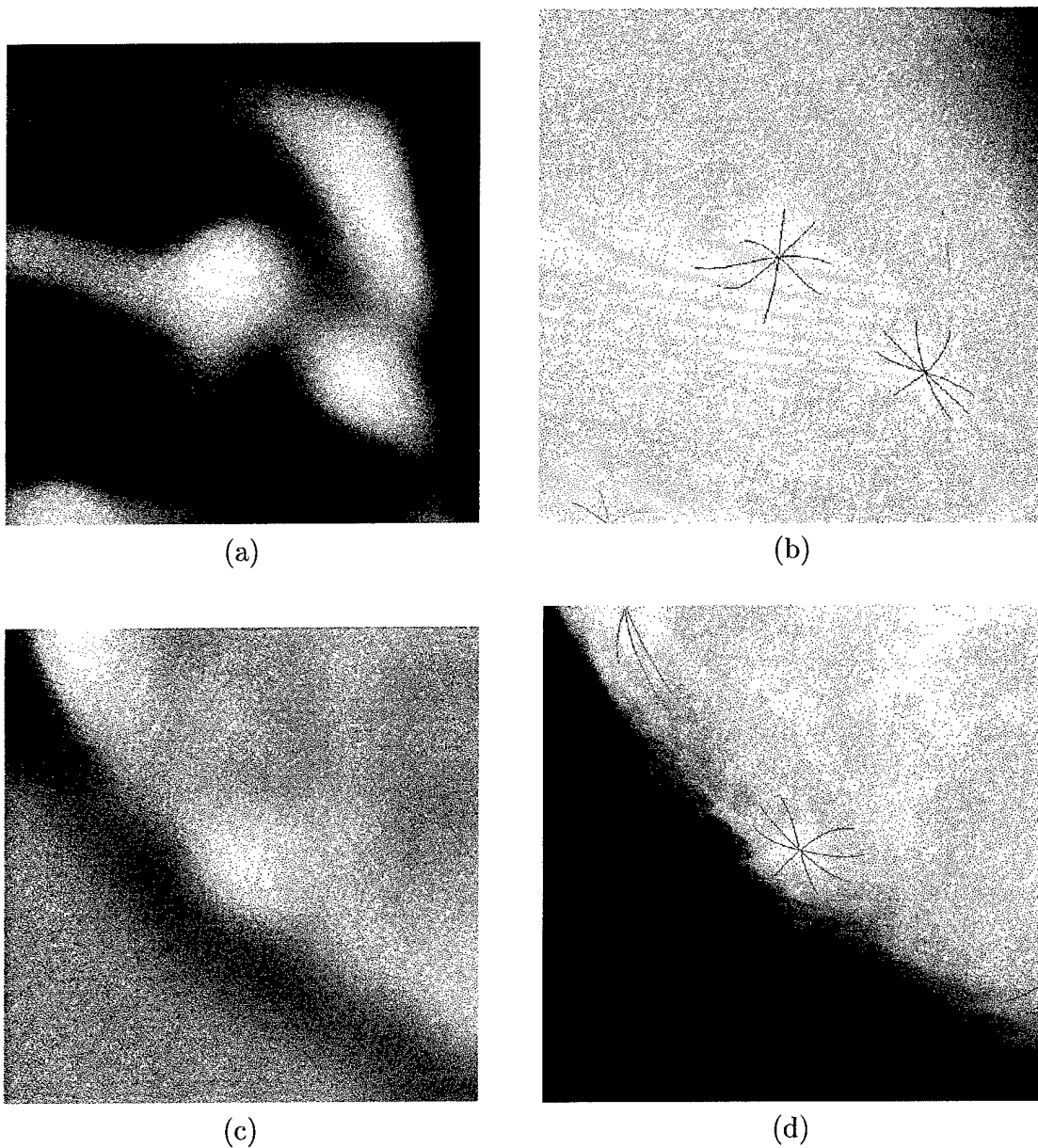
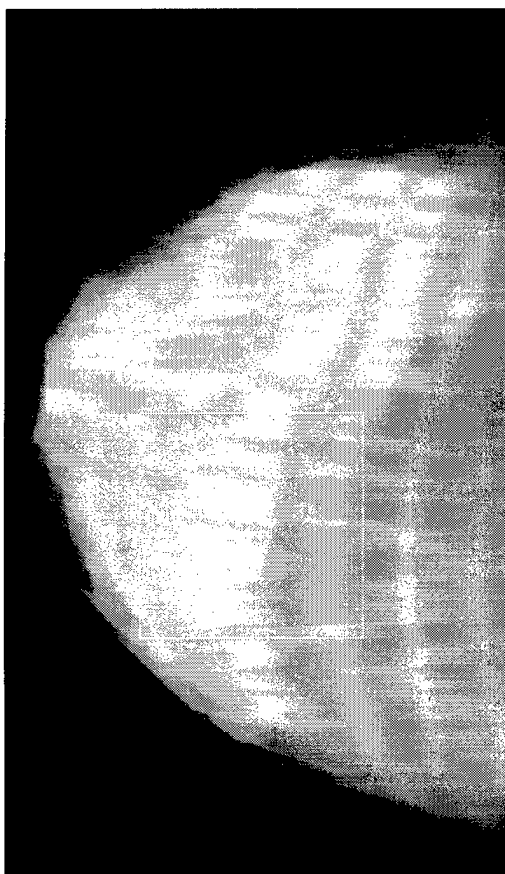
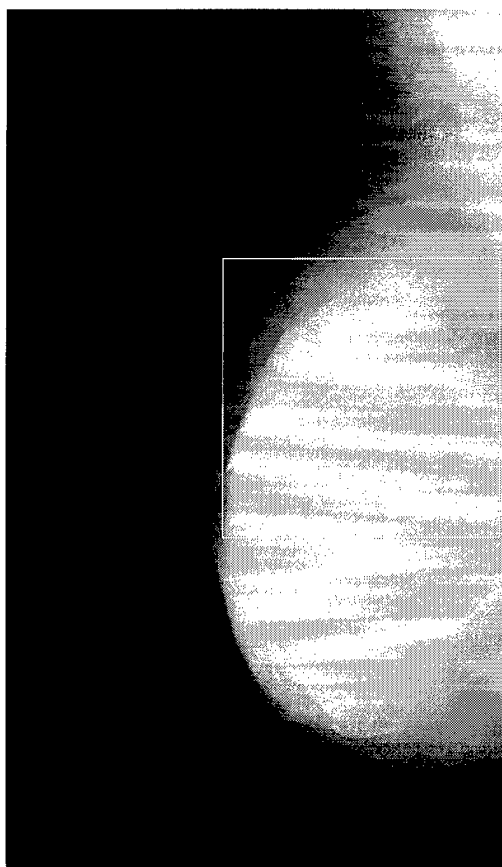


Figure 23: (a) Wavelet coefficients of rml041 at scale 124.8; (b) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass detected in rml041. Note consistency of measure for contours identified within each cluster; (c) Wavelet coefficients of lcc046 at scale 112; (d) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass shown in lcc046.

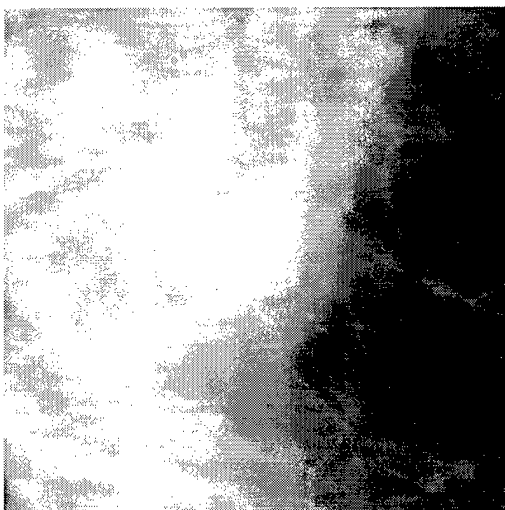




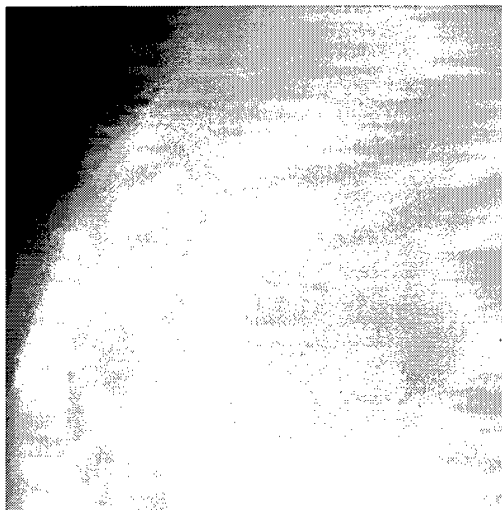
(a)



(b)



(c)



(d)

Figure 24: (a) Mammogram lcc002; (b) Mammogram lml015; (c) ROI cropped from lcc002; (d) ROI cropped from lml015.

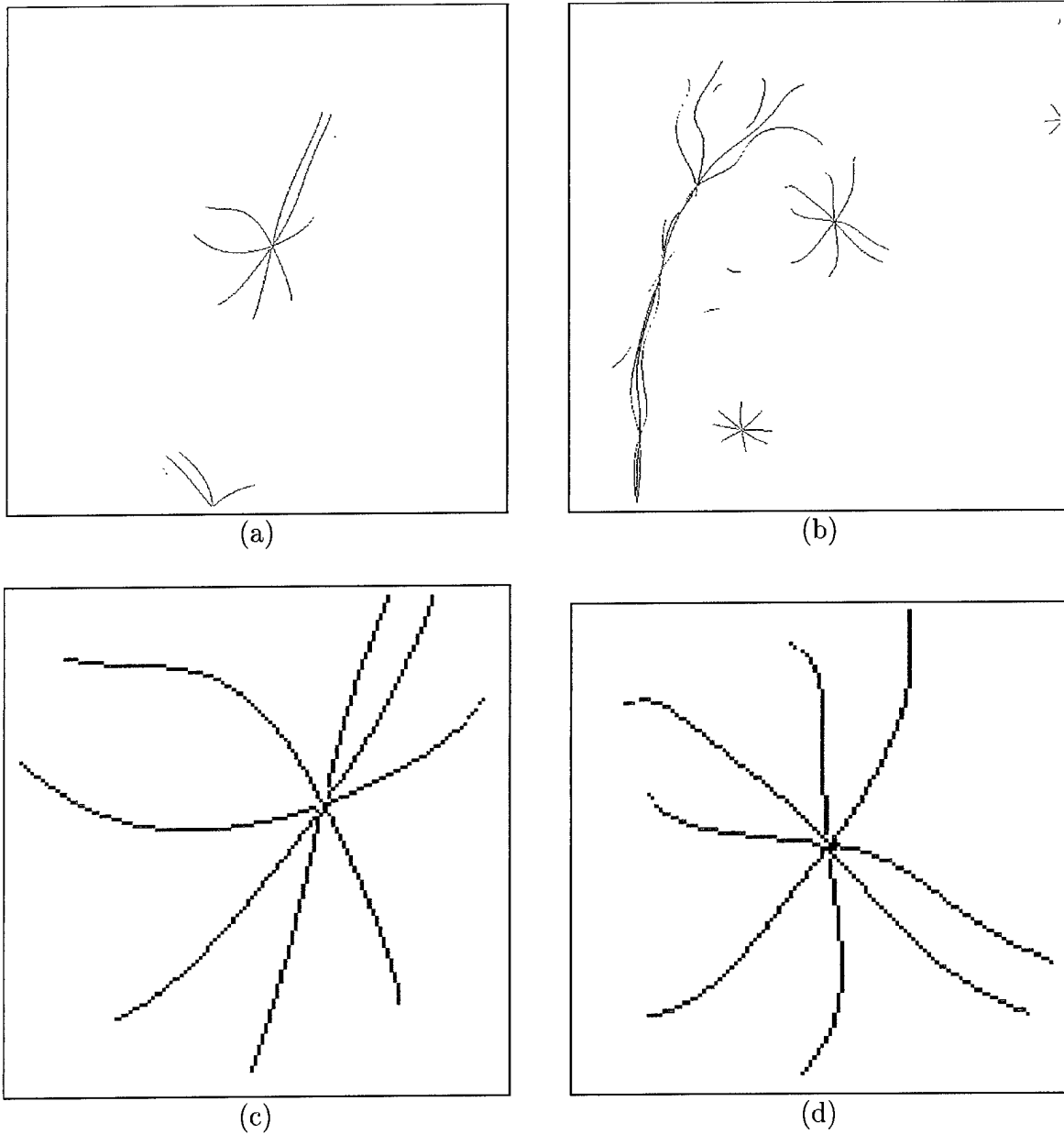
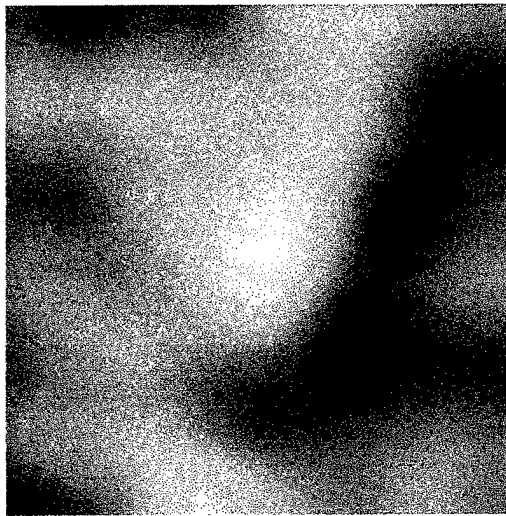


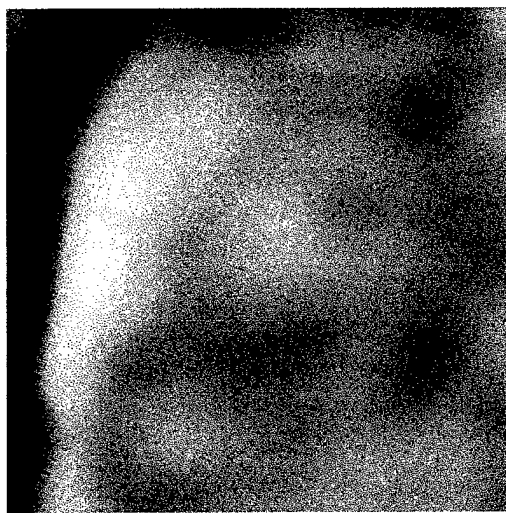
Figure 25: (a) Labeling of wavelet coefficient minima within selected ROI from lcc002 (soft-thresholding applied). “*JUNCTIONs*” are shown in BLUE, “*TERMINALs*” in GREEN, and “*CONTOURs*” in red; (b) Labeling of wavelet coefficient minima within ROI from lml015; (c) 4X magnification of the center cluster representing a subtle mass shown in (a); (d) 4X magnification of the center cluster representing a subtle mass shown in (b).



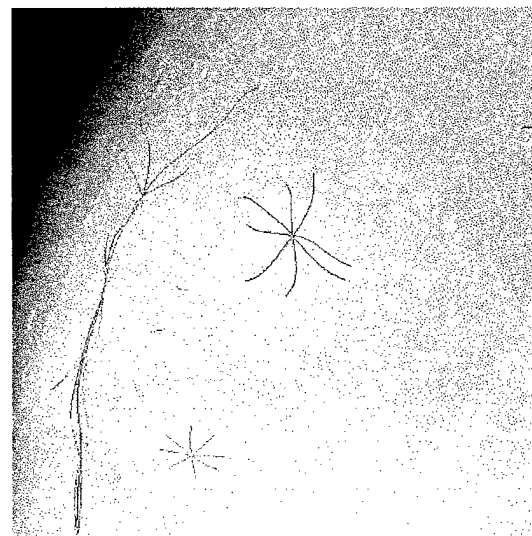
(a)



(b)



(c)



(d)

Figure 26: (a) Wavelet coefficients of lcc002 at scale 126.4; (b) Color coding of a local texture measure computed normal to the contour of a spicular distortion around a mass shown in lcc002. Note consistency of measure for contours identified within each cluster; (c) Wavelet coefficients of lml015 at scale 89.6; (d) Color coding of a local texture measure computed normal to the contour of a spicular distortion around the mass shown in lml015.

a mass. Additional texture information was needed to separated it from the true mass. Cases lcc002 and lml015 are shown in Figure 24. These mammograms were rated as “difficult” cases by an expert mammographer. A selected ROI within each mammogram was decomposed onto an optimum scale 126.4 and 89.6 respectively. Their wavelet minima are shown graphically in Figure 25. Detection results are shown in Figure 26. Both masses were detected. It is important to note that in all cases, the optimum scale found to detect the mass was not power of 2. For comparison with traditional methods, Figure 27 shows failure to detect the mass at dyadic scales. In addition, tests using the RMI phantom

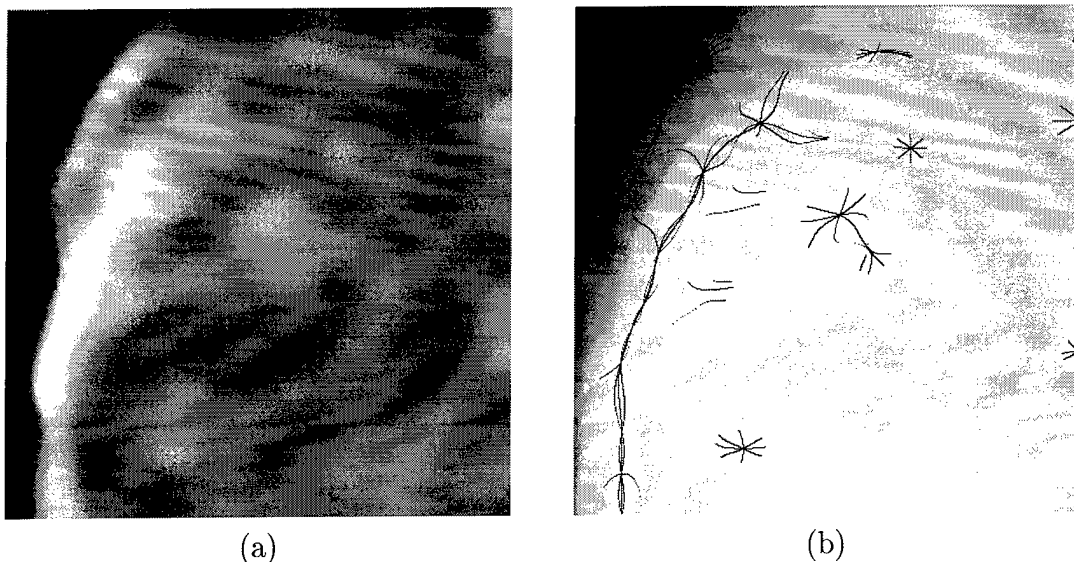


Figure 27: (a) Wavelet coefficients for lml015 at scale 64; (b) Wavelet minima in lml015.

showed that the detection algorithm is quite sensitive. Additional results on real mammograms appear very promising. Masses within dense mammograms of subtle visibility were detected!

### 2.3.7 Frequency Approximation vs Projection Methods

In this section, we discuss the relation between this algorithm with the oblique projection method proposed by Unser *et. al.* etc. [43]. In addition, we show in Theorem 3 that they are equivalent in an extreme case. For a scaling function  $\phi$ , we choose an analysis function<sup>4</sup>  $\phi_1$ . Next the dilated wavelet  $\psi_a(t)$  is projected into  $V(\phi)$  such that the approximation error is orthogonal to  $V(\phi_1)$ . Here

$$p_a(k) = (q_a * q_{12})(k), \quad (85)$$

<sup>4</sup>An analysis function is defined as a function which satisfies only scaling function properties (1) and (2), as described in section 2.3.3 at this chapter.

$$q_a(k) = \langle \psi_a(t), \phi_1(t - k) \rangle, \quad (86)$$

$$a_1(k) = \langle \phi_1(t - k), \phi(t) \rangle, \quad (87)$$

$$q_{12} = a_1^{-1}. \quad (88)$$

When  $\phi_1(t) \in V(\phi)$ ,  $p_a$  is an orthogonal projection. Ideally,  $\phi_1(t)$  should be chosen such that the oblique projection's approximation error is very close to that of an orthogonal projection, and the computation is simple and fast. We chose  $\psi_a(t) = \psi\left(\frac{t}{ss_0}\right)$ , and  $s_0$  is selected based on some estimate of an acceptable projection error. The parameter  $s_0$  sets the finest scale resolution, and the continuous scale discrete wavelet transform of  $f(t)$  is

$$\mathbf{W}_a f(a, n) \approx \sum_{k \in \mathbb{Z}} p_a(k) x(n - k). \quad (89)$$

**Theorem 3** *Let  $\phi(t) = \beta^{n-1}(t)$ ,  $\phi_1(t) = \beta^{m-1}(t)$ . When  $n + m \rightarrow +\infty$ , the oblique projection method is equivalent to the frequency approximation method.*

Please see proof in Appendix A.2. As claimed in Theorem 3, the frequency approximation method is equivalent to the projection method in one limiting case. The advantage of the oblique projection method is that it can be implemented by FIR and IIR filters to achieve  $O(N)$  complexity per scale. However, the frequency based approximation method is easy to expand for multi-dimension signal analysis. Choosing a higher order scaling function does not affect the computation complexity adversely.

### 2.3.8 Summary

We showed that the continuous wavelet transform of a band-limited signal can be estimated from samples of a signal. If a mother wavelet is compact in the frequency domain, effectively no approximation error was observed experimentally.

We described a modified dyadic wavelet transform, which was able to decompose a signal onto an arbitrary scale and reconstruct perfectly the signal from that scale. We presented a special class of spline wavelet filters, which allowed us to use a high order scaling function without significantly increasing the computational complexity of analysis.

The importance of arbitrary scale analysis was demonstrated by digital radiographs of a mammography phantom and digitized mammograms. The study showed that the algorithm was able to detect very subtle masses, which were rated to be almost invisible by radiologist specializing in mammography.

## 2.4 Enhancement of Mammograms from Oriented Information

### 2.4.1 Introduction

Mammograms can detect tumors that are an eighth of an inch in diameter, while manual examination usually fails to detect tumors smaller than a half-inch. Reliable diagnosis by radiographs of malignant breast disease depends on observing local and distant changes in tissues produced by disease. Of the visual signs of cancer found by radiologists, the primary signs of masses and calcifications are most important [54, 55]. Unfortunately, at the early stages of breast cancer, these signs are very subtle and varied in appearance, making diagnosis difficult and challenging even to specialists [54, 56].

The primary radiographic signs of cancer are related to tumor mass and its density, size, shape, borders and calcification content. Extraction of these features and enhancement of them may assist radiologists to locate suspicious areas more reliably [57].

Multiscale representations based on wavelets have been carried out for mammographic feature analysis [32, 58, 59]. Laine *et al.* [32] used two overcomplete multiscale representations for contrast enhancement. Mammograms were reconstructed from transform coefficients modified at each level by nonlinear weighting functions. Qian *et al.* [58] introduced tree-structured nonlinear filters for microcalcification cluster detection. An image was enhanced by tree-structured nonlinear filters with fixed parameters and adaptive order statistic filters. Richardson Jr. [59] applied linear and nonlinear filtering approaches to the analysis of mammograms. Here, a linear multiscale decomposition was obtained via a wavelet transform; a nonlinear multiscale decomposition employed a “mean curvature partial differential equation” filter and “weighted majority-minimum range” filter. In addition, Li *et al.* [60] extended a conventional multiresolution wavelet transform into a multiresolution and multiorientation wavelet transform. They applied directional wavelet analysis to capture orientation information within each mammogram.

Freeman and Adelson [20] first proposed the concept of steerable filters and applied it to several problems in the area of computer vision. With a set of “basis filters”, one can adaptively steer a filter along any orientation. Hilbert transform pairs were constructed to find a local “oriented energy” measure and dominant orientation.

Kass and Witkin [61] developed an algorithm for estimating the orientation of texture patterns. An orientation pattern was decomposed into a flow field, describing the direction of anisotropy, and a residual pattern obtained by describing the image in a coordinate system built from the flow field. The texture orientation was estimated from Laplacian of Gaussian filters. Rao and Schunck [62] proposed another algorithm based on the gradient of the Gaussian. Their new algorithm incorporated a more sophisticated scheme for

computing the coherence of the flow field.

In this chapter, an enhancement algorithm based on multiscale wavelet analysis is described. Features were extracted by separable steerable filters. A coherent image and phase information were then generated. A nonlinear function, integrating coherent image and phase information, was applied to the transform coefficients at each level. An enhanced image was obtained via an inverse wavelet transform of the modified coefficients. The novelty and advantage of this algorithm compared to existing techniques (including those previously developed under this grant) lies in its detection of directional features and removal of unwanted perturbations (artifacts).

## 2.4.2 Background

In this section we briefly describe the mathematical background and fundamental ideas used in subsequent sections.

### Wavelet Transforms

Wavelet transforms have become well recognized as useful tools for many applications in signal processing. A function  $\psi(x)$  is said to be a wavelet if and only if its Fourier transform  $\hat{\psi}(\xi)$  satisfies the *admissibility condition*

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty. \quad (90)$$

This condition implies that

$$\int_{-\infty}^{+\infty} \psi(x) dx = 0. \quad (91)$$

This means that  $\psi(x)$  will have at least some oscillations.

Wavelets constitute a family of functions derived from one single function  $\psi$  (*mother wavelet*) by dilations and translations

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right), \quad (92)$$

where  $a \in R^+, b \in R$ . The idea of the wavelet transform is to represent any function  $f(t)$  as a superposition of wavelets. The continuous wavelet transform is defined as

$$W_f(a,b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx, \quad (93)$$

where  $\bar{\psi}$  denotes the complex conjugate of  $\psi$ . A function can be reconstructed from its wavelet transform by means of the “*resolution of identity*” formula

$$f = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle f, \psi_{a,b} \rangle \psi_{a,b} \frac{dad b}{a^2}. \quad (94)$$

In the practice of digital mammography, one prefers to express  $f$  as a discrete superposition. Let us discretize the translation and dilation parameters of the wavelet in Equation (92):

$$\psi_{m,n}(x) = a_0^{-\frac{m}{2}} \psi(a_0^{-m}x - nb_0), \quad (95)$$

where  $a = a_0^m, b = nb_0 a_0^m$ , with  $m, n \in Z$ , and  $a_0 > 1, b_0 \neq 0$ . On this discrete grid, the wavelet transform is simply

$$W_f(m, n) = a_0^{-\frac{m}{2}} \int_{-\infty}^{+\infty} f(x) \psi(a_0^{-m}x - nb_0) dx. \quad (96)$$

The original signal can be approximated as linear combinations of the wavelet bases,

$$f(x) \approx \sum_{m,n} W_f(m, n) \psi_{m,n}(x). \quad (97)$$

One popular discretization is to choose  $a_0 = 2, b_0 = 1$ ,

$$\psi_{m,n}(x) = 2^{-\frac{m}{2}} \psi(2^{-m}x - n), \quad (98)$$

which are called *dyadic wavelets*.

### Steerable Filters

A function  $f(x, y)$  is called “steerable” if it can be expressed as a linear combination of rotated versions of itself. The fundamental idea of steerable filters is to apply “basis filters” which correspond to a fixed set of orientations and interpolate between each discrete response. Thus, one must decide the number of “basis filters” and the corresponding interpolation functions. As defined in [20] a steering constraint may be formulated by

$$f^\theta(x, y) = \sum_{i=1}^M k_i(\theta) f^{\theta_i}(x, y), \quad (99)$$

where  $M$  is the number of basis functions required to steer a function  $f^{\theta_i}(x, y)$ .

Hereafter, it will be more convenient to work in polar coordinates  $r = \sqrt{x^2 + y^2}$  and  $\phi = \arg(x, y)$ . Let  $f$  be any function that can be expressed as a Fourier series in polar angle  $\phi$ :

$$f(r, \phi) = \sum_{n=-N}^N a_n(r) e^{jn\phi}, \quad (100)$$



where  $j = \sqrt{-1}$ .

The theorem below was posed by [20] and is included for the clarity of discussion.

**Theorem 1:** The steering condition (99) holds for a function  $f$  expanded in the form of (100) *if and only if* the interpolations  $k_i(\theta)$  are solutions of

$$\begin{pmatrix} 1 \\ e^{j\theta} \\ \vdots \\ e^{jN\theta} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \cdots & e^{j\theta_M} \\ \vdots & \vdots & \ddots & \vdots \\ e^{jN\theta_1} & e^{jN\theta_2} & \cdots & e^{jN\theta_M} \end{pmatrix} \begin{pmatrix} k_1(\theta) \\ k_2(\theta) \\ \vdots \\ k_M(\theta) \end{pmatrix}. \quad (101)$$

Then,  $f^\theta(r, \phi)$  may be expressed as

$$f^\theta(r, \phi) = \sum_{i=1}^M k_i(\theta) g_i(r, \phi), \quad (102)$$

where  $g_i(r, \phi)$  can be any set of functions.

## Measure of Coherence

Texture plays an important role in many machine vision and image processing tasks including surface inspection, scene classification, surface orientation and shape determination [63]. Texture patterns may be characterized by extracting measurements that quantify the nature and directions of pattern. Most breast carcinomas have the appearance of stellate lesions consisting of a central mass surrounded by radiating spicules [54]. The spicules radiate outward in all directions and vary in length. This provides an important cue for early cancer detection.

Much attention has been given to the notion of decomposing an intensity image into intrinsic images to extract meaningful information [64, 65]. These intrinsic properties represent basic components of the image formation process and therefore reveal features “hidden” inside an image. The information they provide is beyond the intensity image alone.

Rao and Schunck [62] defined the orientation field of a texture image to consist of two images — an angle image and a coherence image. The angle image denotes the dominant local orientation at each point and the coherence image represents the degree of anisotropy at each point. They strongly advocated the use of angle and coherence images as intrinsic images. In this chapter, we investigated the efficiency of these two representations to capture and enhance features of importance to mammography.

### 2.4.3 Methodology

Our algorithm consists of the following four steps.

**(1) Multiscale Wavelet Transform:** Wavelet transforms, owing to their localization characteristics, are powerful tools of analysis for many signal and image processing applications. Multiscale analysis can extract features at distinct scales and provide local information often hidden in an original mammogram. In our algorithm, a digitized mammogram was decomposed using a fast wavelet transform algorithm (FWT) [7]. In order to obtain wavelet coefficients at each level without downsampling, an “*algorithme à trous*” (algorithm with holes) [13, 66] was implemented. Let  $s^{i+1}$  denote an original image and  $D^i f$  be obtained by inserting  $2^i - 1$  zeros between every pair of the coefficients representing  $f$ .  $(D^i f)_x$  and  $(D^i f)_y$  stand for carrying out convolution operations with the filter  $D^i f$  along  $x$  and  $y$  directions, respectively. The decomposition and reconstruction equations at level  $i$  are as follows:

Decomposition:

$$\begin{aligned} s^{i+1} &= s^i * (D^i h)_x * (D^i h)_y, \\ w_x^{i+1} &= s^i * (D^i g)_x, \\ w_y^{i+1} &= s^i * (D^i g)_y. \end{aligned} \quad (103)$$

Reconstruction:

$$s^i = w_x^{i+1} * (D^i k)_x * (D^i l)_y + w_y^{i+1} * (D^i l)_x * (D^i k)_y + s^{i+1} * (D^i \tilde{h})_x * (D^i \tilde{h})_y, \quad (104)$$

where “ $*$ ” indicates discrete convolution, and  $h, g, k$ , and  $l$  are filters whose Fourier transforms  $(H(\omega), G(\omega), K(\omega)$ , and  $L(\omega)$ , respectively) satisfy [7]

$$\begin{aligned} G(\omega)K(\omega) + |H(\omega)|^2 &= 1, \\ L(\omega) &= \frac{1 + |H(\omega)|^2}{2}. \end{aligned} \quad (105)$$

**(2) Separable Steerable Filters:** A filter is called “steerable” if the filter at an arbitrary orientation can be expressed as a linear combination of a set of basis filters, generated from rotations of a single kernel [20]. Steerable filters [20, 67, 68], which can be adaptively adjusted to arbitrary orientation, were used to detect stellate patterns of spicules and locate feature orientations more precisely. As pointed out by [21], the separability property of the filters sped up computations considerably when convolved with a large image matrix. In our algorithm, we used three basis functions as steerable filters. The  $x$ - $y$  separable steerable approximations of filter kernels were generated by Singular Value Decomposition (SVD) [20, 21]. Using a set of separable steerable filters, local energy ( $M^i$ ) and associated dominant directions ( $A^i$ ) were determined by the basis functions and their Hilbert

transforms [20, 69, 70, 71]. We applied separable steerable filters to an original image  $S^0$  and low-pass filtered images at each level  $S^i, i = 1, \dots, n$ , to capture salient multiscale features. Discrete realizations of the three basis functions applied in our algorithm are shown in Figure 28.

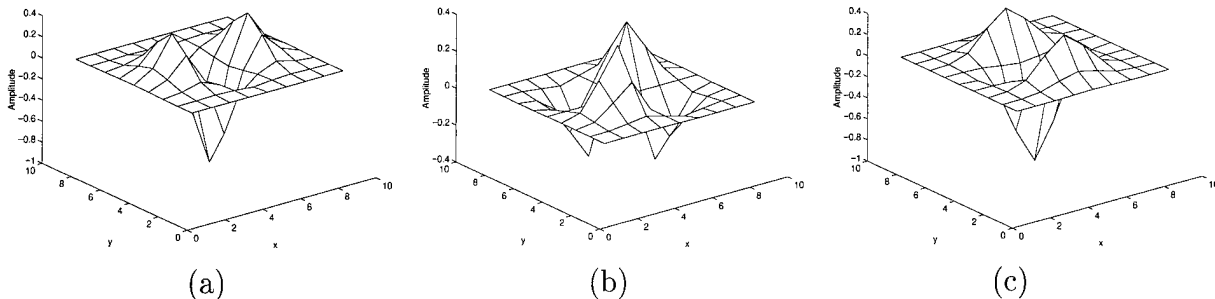


Figure 28: Two-dimensional impulse responses of three basis functions.

**(3) Coherence Maps:** A coherence map is an image showing a local measure of the degree of anisotropy of flow [61, 62]. If the orientations of a texture pattern at any point  $(x_i, y_i)$  are coherent, then magnitude and phase information are important and should be emphasized. Conversely, if the orientations are not coherent, the magnitude and phase information can be neglected or attenuated. Kass and Wilkin [61] suggested a simple way of measuring strength of coherence by finding the ratio

$$\chi(j, k) = \frac{|W(j, k)J(j, k)|}{(W(j, k)|J(j, k)|)}, \quad (106)$$

where  $W(j, k)$  denotes a local weighting function with unit integral,  $J(j, k)$  denotes the squared gradient vector at  $(j, k)$ , and  $|\cdot|$  denotes absolute value.

An alternative measure of coherence was proposed by Rao and Schunck [62] and was obtained by weighting the energy with the normalized projection of energy within a specified window ( $\mathcal{W}$ ) onto the central point  $(j, k)$  of the window. The coherence ( $C^i$ ) was expressed as

$$C^i(j, k) = M^i(j, k) \frac{\sum_{(m,n) \in \mathcal{W}} |M^i(m, n) \cos(A^i(j, k) - A^i(m, n))|}{\sum_{(m,n) \in \mathcal{W}} M^i(m, n)}, \quad (107)$$

where  $M^i(j, k)$  and  $A^i(j, k)$  denote energy and phase of point  $(j, k)$  at level  $i$ , respectively. This coherence measure incorporated the gradient magnitude and hence placed more

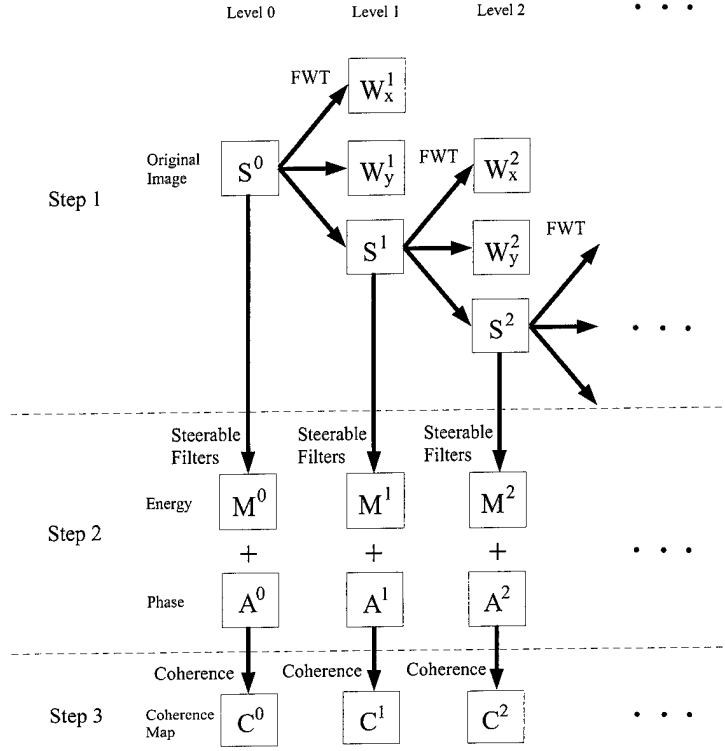


Figure 29: Overview of processing for Steps 1–3.

weight on regions that had higher visual contrast. This method performed better than previous measures applied to the similar data [62]. In our algorithm, we implemented this method to measure the coherence map at each level. A  $5 \times 5$  window was used to perform smoothing. The measure of coherence  $C^i$  was obtained from Equation (107). The coherence and orientation data extracted the more salient features of spiculated lesions. A schematic diagram of Steps 1–3 is shown in Figure 29.

**(4) Nonlinear Operators:** So far, we have computed all the information we need in our algorithm. A nonlinear operation was then applied within each level to precisely modify transform coefficients. This operation integrated both coherence map and phase information.

**A. Modification from Coherence Map:** Let  $C^i(j, k)$  denote the coherence measure of point  $(j, k)$  at some level  $i$ . Modifications from coherence measure were obtained by a nonlinear function expressed as

$$\begin{aligned} MC^i(j, k) &= e^{\log_{10} \sqrt{C^i(j, k)}} \quad , \text{ if } C^i(j, k) \neq 0, \\ &= 0 \quad , \text{ if } C^i(j, k) = 0, \end{aligned}$$

such that a coefficient was emphasized if its coherence measure was large and attenuated if

small.

**B. Modification from Phase:** Phase information is important to characterize well an oriented texture and therefore we did not neglect its contribution to the modification of coefficients. We applied a sinusoidal weighting to the phase information. The “detail” sub-bands of wavelet coefficients obtained in Step 1 included two components: the component along  $x$  direction and the component along  $y$  direction. The  $x$  component was obtained by high-pass filtering along  $x$  direction, hence mostly vertical features within the mammogram were detected. We emphasized the points whose dominant orientations were near 0 and  $\pi$  (with respect to the vertical axis). Thus, the modification from phase information was

$$MA^i(j, k) = 0.01 + |\cos(A^i(j, k))|. \quad (108)$$

Note that a constant (0.01) was added to the above modification so that the phase factor would not be neglected when the phase was equal to  $\frac{\pi}{2}$  or  $\frac{3\pi}{2}$ .

The  $y$  component was obtained by high-pass filtering along the  $y$  direction, hence horizontal features of an image were detected. We emphasized the points whose dominant orientations were near  $\frac{\pi}{2}$  and  $\frac{3\pi}{2}$  (with respect to the vertical axis). The modification from phase information was

$$MA^i(j, k) = 0.01 + |\sin(A^i(j, k))|. \quad (109)$$

Note that a constant (0.01) was added to the above modification so that the phase factor would not be neglected when the phase was equal to 0 or  $\pi$ .

The modifications from coherence map and phase at level  $i - 1$  were combined to adjust the wavelet coefficients at level  $i$ . The final modification was therefore

$$MW^i = T^{i-1} \cdot MC^{i-1} \cdot MA^{i-1} \cdot W^i, \quad (110)$$

where  $T^i$  was a constant at each level. A schematic diagram of the nonlinear operator at level  $i$  is shown in Figure 30. These modified coefficients were then reconstructed, via an inverse fast wavelet transform, to enhance the visualization of possible lesions.

#### 2.4.4 Experimental Results

Our algorithm was applied to enhance mammograms using oriented information. In order to capture distinct directions of subtle features, steerable filters were used. Figure 31 shows the capability of steerable filters for detecting and enhancing objects of distinct direction and scale. The phantom image contained 24 diamond shaped objects with four directions

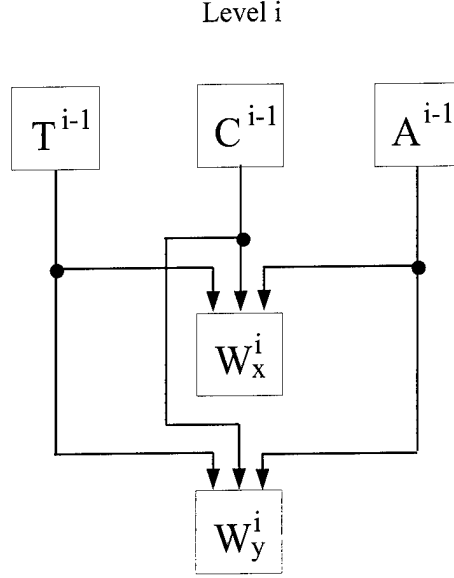


Figure 30: A schematic diagram of nonlinear operator at level  $i$ .

( $0^\circ$ ,  $30^\circ$ ,  $60^\circ$ , and  $90^\circ$ ) and six scales, as shown in Figure 31(a). After applying our algorithm to this image, the borders of these objects at each scale were clearly enhanced. The phantom objects in the last two rows of the image, were very close to microcalcifications. This demonstrated the ability of this algorithm to detect microcalcification clusters in mammograms, without artifacts.

Three examples of malignant lesions with distinct radiographic signs of cancer were processed to show the effectiveness of our algorithm. The images were of matrix size ( $512 \times 512$ ). For each case, both global and regions of interest (ROI) were shown along with the corresponding enhanced ROI image.

### Calcifications

Figure 32 shows a mammogram (mam009lml) with microcalcification clusters. The original mammogram is shown in Figure 32(a). Figure 32(b) shows an original suspicious area of the mammogram. After enhancement, clusters of calcifications appear clearly in the center of the image.

### Stellate lesions

A mammogram (mam041rcc) with a stellate lesion is shown in Figure 33. Figure 33(a) shows the original image. Figure 33(b) presents an original digital radiograph with a partially obscured irregular mass in the center of the image matrix. After applying our algorithm to this image, the enhanced image makes obvious spiculated lesions around the mass. The mass itself was also enhanced, as shown in Figure 33(c). These clear spiculated patterns suggested radiologists that this mass is more likely malignant, rather than benign.

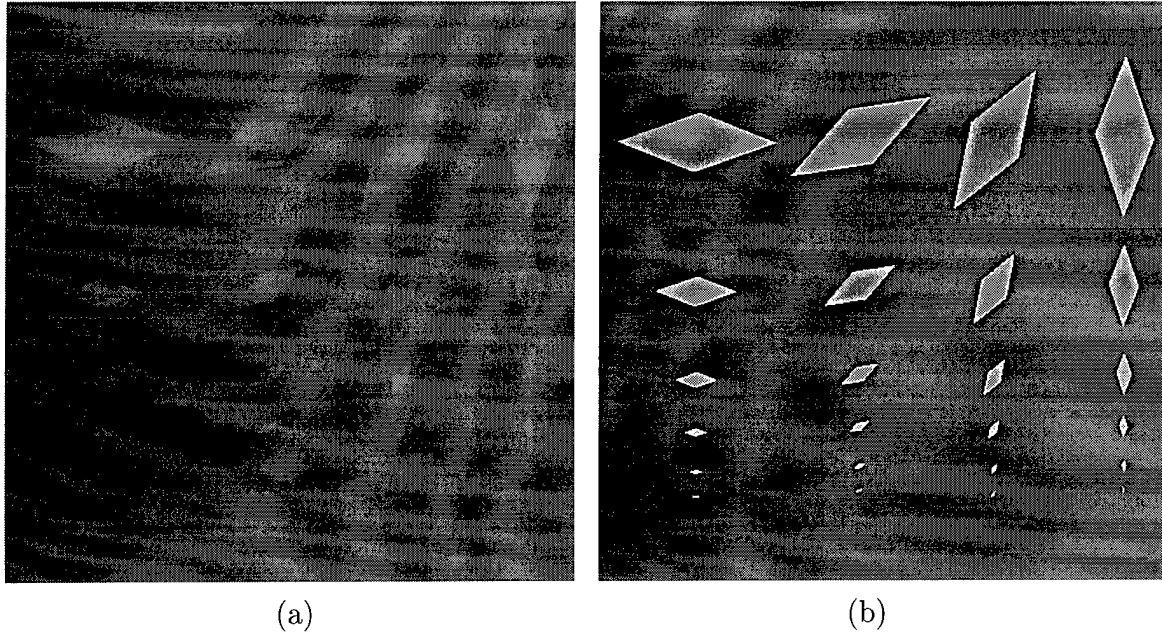


Figure 31: The ability of detecting objects oriented along distinct directions at various scales: (a) Original test image (mathematical phantom), (b) enhanced image after processing.

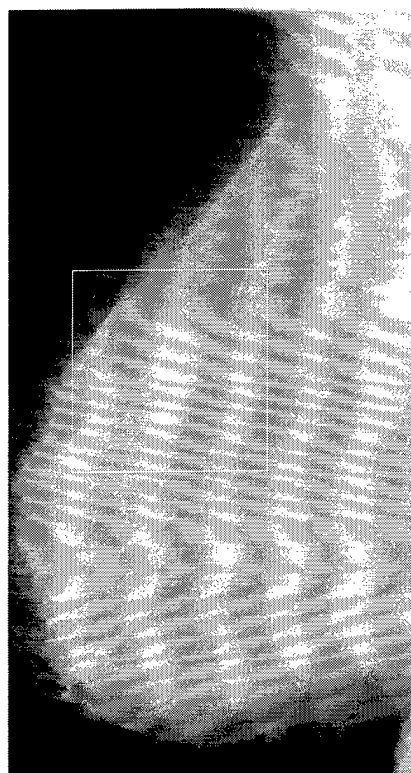
## Masses

A mammogram (mam004lcc) with a mass tumor is shown in Figure 34. The craniocaudal view of the left breast shown in Figure 34(b) shows an irregular spiculated mass in retroglandular fat. The enhanced version shown in Figure 34(c) better delineates the margins of the mass.

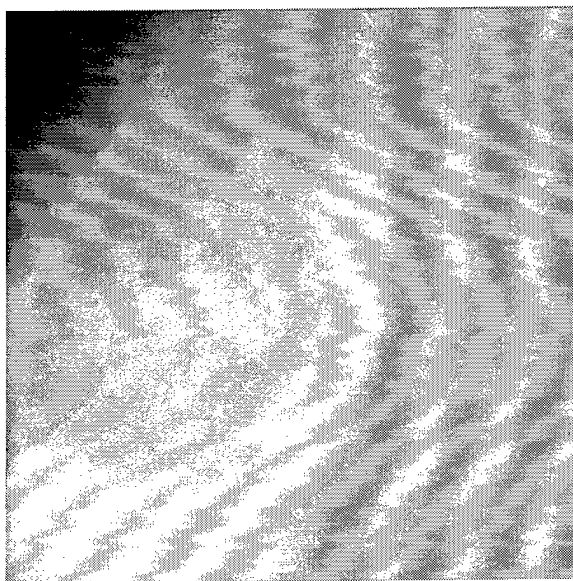
### 2.4.5 Summary

An enhancement algorithm relying on multiscale wavelet analysis and extracting oriented information at each scale of analysis was investigated. The evolution of wavelet coefficients across scales characterized the local shape of irregular structures. Using oriented information to detect the features of an image appears to be a promising approach for enhancing complex and subtle structures of the breast. Steerable filters which can be rotated at arbitrary orientations can reliably find visual cues within each spatial-frequency sub-band of an image. “Coherence measure” and “dominant orientation” clearly helped us to discriminate features from complex surrounding tissue typical in mammograms.

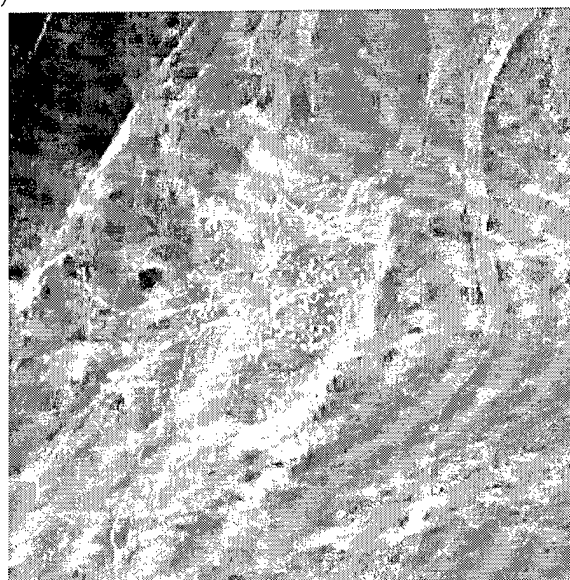
Existing and previous multiscale enhancement approaches [72, 32, 59] attempted to enhance an image by detecting edges. Unfortunately, most edge detection algorithms can not distinguish between “authentic” edges and phantom edges. In contrast, this algorithm relied upon a coherence map and phase information which resulted in an enhancement



(a)



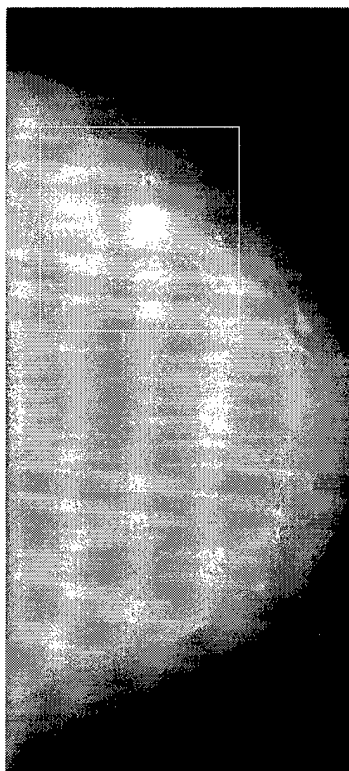
(b)



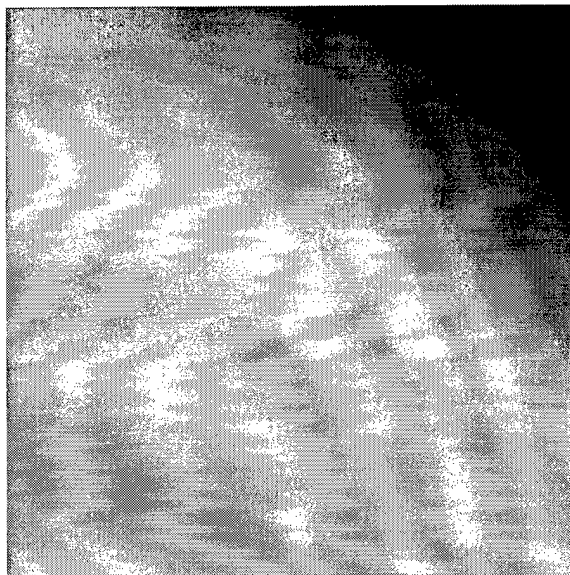
(c)

Figure 32: Mammogram with calcifications: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image.

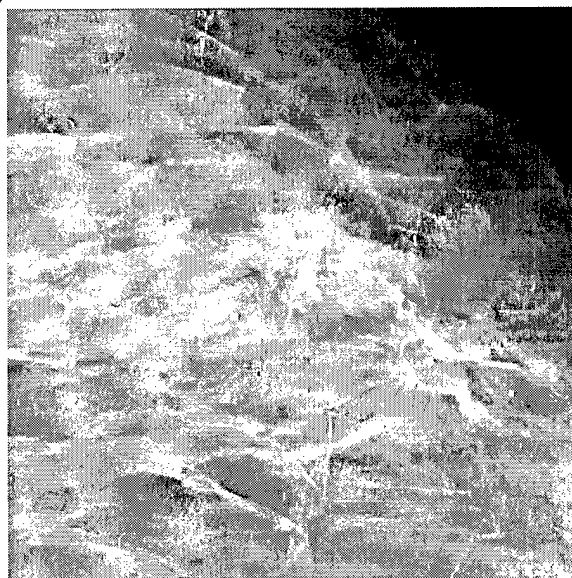




(a)

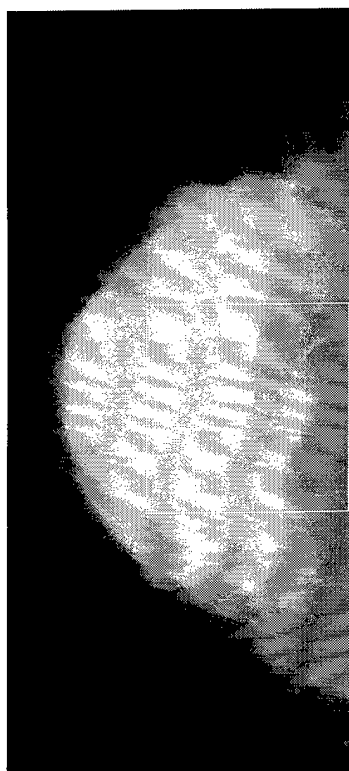


(b)

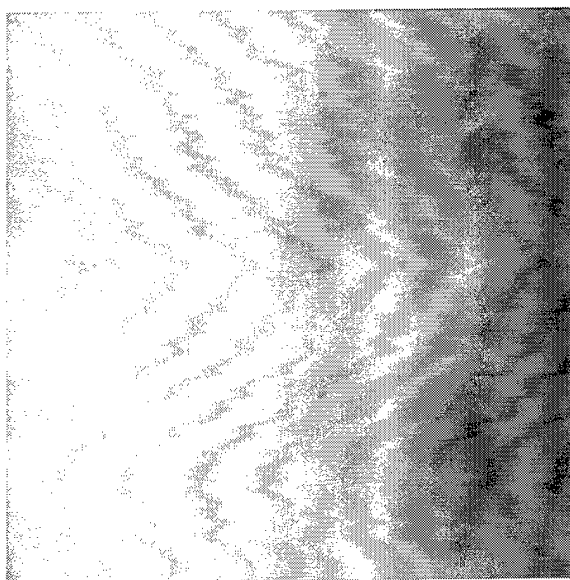


(c)

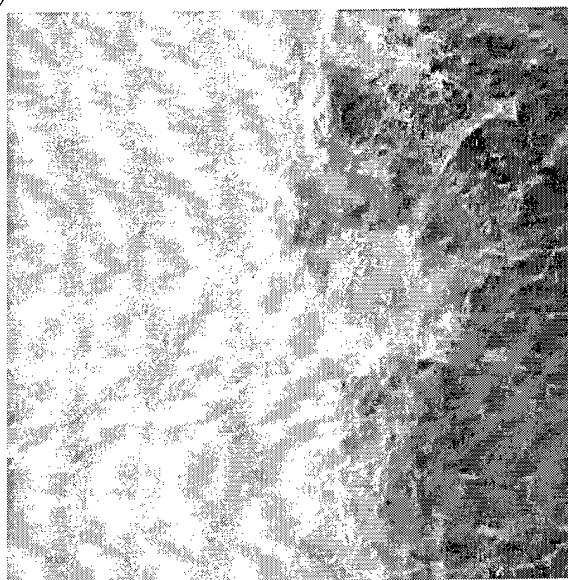
Figure 33: Mammogram with a stellate lesion: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image.



(a)



(b)



(c)

Figure 34: Mammogram with a mass: (a) Original mammogram, (b) ROI image, (c) enhanced ROI image.

naturally close to the original image. This type of artifact free enhanced image can provide more “obvious” (and familiar) visual cues for radiologists.

## 2.5 Multivoice Undecimated Wavelet Transforms

### 2.5.1 Introduction

For many analysis/synthesis applications the sampling provided by the traditional wavelet transform is inadequate. In particular, a finer grid is needed when computing the wavelet transform for applications that require “good” time–frequency localization. This can be achieved by computing an undecimated wavelet transform and the addition of voices. In this chapter we review the tools that will allow us to compute an exact multivoice undecimated wavelet transform for an arbitrary wavelet. Furthermore, we introduce a wavelet function that exhibits nearly optimum time–frequency localization, a desirable property in signal analysis as we will see below.

The chapter is organized as follows: In Sections 2.5.2 and 2.5.3 we review the short–time Fourier transform and the wavelet transform, including the theory of frames, which will allow us to carry out wavelet analysis/synthesis for an arbitrary wavelet. In Sections 2.5.4 and 2.5.5 we review the à trous and Mallat’s algorithms and learn that although the à trous algorithm was originally devised as a computationally efficient implementation of the wavelet transform, it is more properly viewed as a nonorthogonal multiresolution decomposition for which the discrete wavelet transform computes a sampled continuous wavelet transform exactly. In Sections 2.5.6 and 2.5.7 we introduce the sine-Gabor wavelet and show that it exhibits nearly optimum time–frequency localization and satisfies the conditions of a wavelet frame. Finally, in Section 2.5.8, we summarize our results and describe related ongoing work.

### 2.5.2 The Short-Time Fourier Transform

The goal of signal analysis is to extract relevant information from a signal by transforming it into a representation where the properties of the signal are more evident. For the analysis of stationary signals, that is, signals whose properties do not evolve with time, one example of such representation is the Fourier transform

$$(\mathcal{F}s)(\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt.$$

The Fourier transform can be viewed as the inner product of the signal  $s(t)$  and the sinusoidal wave  $e^{j\omega t}$ . The analysis coefficient  $\hat{s}(\omega) = (\mathcal{F}s)(\omega)$  measures the “strength” of the sinusoidal wave of frequency  $\omega$  in the signal  $s(t)$ .

Fourier analysis works well for signals composed of a few stationary components. However, any abrupt change in time in a non-stationary signal  $s(t)$  is spread out over the whole frequency axis in  $\hat{s}(\omega)$ . In order to adapt the Fourier transform to non-stationary signals, Gabor [73] introduced a new transform by using a windowed function  $w(t)$  in the Fourier integral

$$(\mathcal{G}_w s)(\tau, \omega) = \int_{-\infty}^{\infty} s(t) \overline{w(t - \tau)} e^{-j\omega t} dt.$$

A function  $w(t)$  qualifies as a window function if it is possible to identify its center and standard deviation (or root mean square (RMS) duration) defined by

$$m(w) = \frac{1}{\|w\|_2^2} \int_{-\infty}^{\infty} t |w(t)|^2 dt$$

and

$$\sigma(w) = \frac{1}{\|w\|_2} \left\{ \int_{-\infty}^{\infty} (t - m(w))^2 |w(t)|^2 dt \right\}^{1/2}, \quad (111)$$

respectively [74]. In the original Gabor transform, the window function was a Gaussian [73], however, the transform is valid for any type of window function and is traditionally referred to as the short-time or window Fourier transform.

The short-time Fourier transform can be viewed as the inner product of the signal  $s(t)$  with the family of functions

$$w_{\tau, \omega}(t) = w(t - \tau) e^{j\omega t}, \quad (\tau, \omega) \in \mathbb{R}^2,$$

that is,

$$(\mathcal{G}_w s)(\tau, \omega) = \langle s(t), w_{\tau, \omega}(t) \rangle. \quad (112)$$

It is easy to show that  $w_{\tau, \omega}(t)$  is a window function with center and standard deviation given by

$$m(w_{\tau, \omega}) = m(w) + \tau \quad (113)$$

and

$$\sigma(w_{\tau, \omega}) = \sigma(w), \quad (114)$$

respectively. From Equations 112, 113 and 114 we see that in the time domain the analysis coefficient  $(\mathcal{G}_w s)(\tau, \omega)$  essentially depends on the values of  $s(t)$  for  $t \in [m(w) + \tau - \sigma(w), m(w) + \tau + \sigma(w)]$ . This is called time localization. It follows from

the above observation that two pulses in time can be discriminated only if they are more than  $2\sigma(w)$  apart. This is referred to as the time resolution of the short-time Fourier transform.

By applying Parseval's theorem to Equation 112 we obtain an alternative formula for the short-time Fourier transform

$$(\mathcal{G}_w s)(\tau, \omega) = \frac{1}{2\pi} \langle \hat{s}(\gamma), \hat{w}_{\tau, \omega}(\gamma) \rangle, \quad (115)$$

where

$$\hat{w}_{\tau, \omega}(\gamma) = \hat{w}(\gamma - \omega) e^{-j(\gamma - \omega)\tau}, \quad (\tau, \omega) \in \mathbb{R}^2.$$

Suppose  $\hat{w}(\gamma) = (\mathcal{F}w)(\gamma)$  is also a window function with center,  $m(\hat{w})$ , and standard deviation,  $\sigma(\hat{w})$ . Then it is easy to show that  $\hat{w}_{\tau, \omega}(\gamma)$  is a window function with center and standard deviation given by

$$m(\hat{w}_{\tau, \omega}) = m(\hat{w}) + \omega \quad (116)$$

and

$$\sigma(\hat{w}_{\tau, \omega}) = \sigma(\hat{w}), \quad (117)$$

respectively. From Equations 115, 116 and 117 we see that in the frequency domain the analysis coefficient  $(\mathcal{G}_w s)(\tau, \omega)$  essentially depends on the values of  $\hat{s}(\gamma)$  for  $\gamma \in [m(\hat{w}) + \omega - \sigma(\hat{w}), m(\hat{w}) + \omega + \sigma(\hat{w})]$ . This is called frequency localization. It follows from the above observation that two pure sinusoids can be discriminated only if their frequencies are more than  $2\sigma(\hat{w})$  apart. This is referred to as the frequency resolution of the short-time Fourier transform.

From Equations 112 through 117 we observe that  $(\mathcal{G}_w s)(\tau, \omega)$  yields a time-frequency representation of  $s(t)$ . The analysis coefficient  $(\mathcal{G}_w s)(\tau, \omega)$  depends on the time-frequency window

$$[m(w) + \tau - \sigma(w), m(w) + \tau + \sigma(w)] \times [m(\hat{w}) + \omega - \sigma(\hat{w}), m(\hat{w}) + \omega + \sigma(\hat{w})].$$

Increasingly accurate localization in time and frequency is not possible because the area of the time-frequency window is lower bounded by 2, that is

$$4\sigma(w)\sigma(\hat{w}) \geq 2. \quad (118)$$

This is referred to as the uncertainty principle or Heisenberg inequality. Equality is satisfied if and only if the window function is a Gaussian as in the Gabor transform.

The short-time Fourier transform is an isometry (to a proportionality coefficient) from  $L^2(\mathbb{R})$  into  $L^2(\mathbb{R}^2)$  [75], that is

$$\|s\|_2^2 = \frac{1}{2\pi\|w\|_2^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |(\mathcal{G}_w s)(\tau, \omega)|^2 d\tau d\omega.$$

The function  $s(t)$  is reconstructed from  $(\mathcal{G}_w s)(\tau, \omega)$  with the formula

$$s(t) = \frac{1}{2\pi\|w\|_2^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\mathcal{G}_w s)(\tau, \omega) w(t - \tau) e^{j\omega t} d\tau d\omega.$$

The short-time Fourier transform is a redundant representation. Instead of computing  $(\mathcal{G}_w s)(\tau, \omega)$  for all values  $(\tau, \omega)$  in  $\mathbb{R}^2$ , it is possible to uniformly sample both  $\tau$  and  $\omega$  such that the representation is complete and stable [75]. Let  $\tau_0$  and  $\omega_0$  be the sampling intervals in the time and frequency domains respectively. Then, the discrete short-time Fourier transform is defined by [75]

$$(\mathcal{G}_w s)_{m,n} = (\mathcal{G}_w s)(m\tau_0, n\omega_0), \quad (m, n) \in \mathbb{Z}^2.$$

To reconstruct any function  $s(t) \in L^2(\mathbb{R})$  from the set of samples  $(\mathcal{G}_w s)_{m,n}$ ,  $(m, n) \in \mathbb{Z}^2$ , the operator

$$L^2(\mathbb{R}) \xrightarrow{\mathcal{G}_w} \ell^2(\mathbb{Z}^2)$$

must be invertible on its range and have a bounded inverse [75]. In order to invert  $\mathcal{G}_w$ , Daubechies [76] has shown that  $\tau_0$  and  $\omega_0$  must verify

$$\omega_0 \tau_0 < 2\pi.$$

A drawback of the short-time Fourier transform is that once a window function has been chosen the time-frequency resolution of the transform is fixed over the entire time-frequency plane. It is therefore impossible with the short-time Fourier transform to analyze at the same time transient signal components with good time resolution and quasi-stationary signal components with good frequency resolution [77].

### 2.5.3 The Wavelet Transform

The wavelet transform overcomes the resolution limitation of the short-time Fourier transform by letting the time and frequency resolution vary in the time-frequency plane. The wavelet transform is defined by [74]

$$(\mathcal{W}_\psi s)(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} s(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt, \quad s(t) \in L^2(\mathbb{R}), \quad a, b \in \mathbb{R}, \quad (a \neq 0). \quad (119)$$

In order to reconstruct  $s(t)$  from its wavelet transform, the Fourier transform of  $\psi(t)$  must satisfy [74]

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\gamma)|^2}{|\gamma|} d\gamma < \infty. \quad (120)$$

A function  $\psi(t)$  is said to be a basic wavelet if it is possible to reconstruct  $s(t)$  from its wavelet transform.

For any  $\psi(t)$  satisfying Equation 120, the wavelet transform is an isometry (to a proportionality coefficient) from  $L^2(\mathbb{R})$  into  $L^2(\mathbb{R}^2)$  [74], that is

$$\|s\|_2^2 = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |(\mathcal{W}_\psi s)(a, b)|^2 \frac{da}{a^2} db.$$

The function  $s(t)$  is reconstructed from  $(\mathcal{W}_\psi s)(a, b)$ ,  $a, b \in \mathbb{R}$ , ( $a \neq 0$ ), with the formula [74]

$$s(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\mathcal{W}_\psi s)(a, b) \left\{ |a|^{-\frac{1}{2}} \psi \left( \frac{t-b}{a} \right) \right\} \frac{da}{a^2} db.$$

The wavelet transform can be viewed as the inner product of the signal  $s(t)$  with the family of functions

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi \left( \frac{t-b}{a} \right), \quad a, b \in \mathbb{R}, (a \neq 0), \quad (121)$$

that is,

$$(\mathcal{W}_\psi s)(a, b) = \langle s(t), \psi_{a,b}(t) \rangle. \quad (122)$$

Suppose  $\psi(t)$  is a window function with center  $m(\psi)$  and standard deviation  $\sigma(\psi)$ . Then, it is easy to show that  $\psi_{a,b}(t)$  is a window function with center and standard deviation given by

$$m(\psi_{a,b}) = am(\psi) + b \quad (123)$$

and

$$\sigma(\psi_{a,b}) = |a|\sigma(\psi), \quad (124)$$

respectively. From Equations 122, 123 and 124 we have that in the time domain the analysis coefficient  $(\mathcal{W}_\psi s)(a, b)$  essentially depends on the values of  $s(t)$  for  $t \in [am(\psi) + b - |a|\sigma(\psi), am(\psi) + b + |a|\sigma(\psi)]$ . It can be verified [74] that if a basic wavelet,  $\psi(t)$ , is also a window function then  $\psi(t)$  is necessarily in  $L^1(\mathbb{R})$  so that its Fourier

transform,  $\hat{\psi}(\omega)$  is a continuous function. It follows from Equation 120 that  $\hat{\psi}(\omega)$  must vanish at the origin, that is,

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (125)$$

By applying Parseval's theorem to Equation 122 we obtain an alternative formula for the wavelet transform

$$(\mathcal{W}_{\psi}s)(a, b) = \frac{1}{2\pi} \left\langle \hat{s}(\gamma), \hat{\psi}_{a;b}(\gamma) \right\rangle, \quad (126)$$

where

$$\hat{\psi}_{a;b}(\gamma) = |a|^{\frac{1}{2}} \hat{\psi}(a\gamma) e^{-j\gamma b}, \quad a, b \in \mathbb{R}, \quad (a \neq 0).$$

Suppose  $\hat{\psi}(\gamma)$  is also a window function with center,  $m(\hat{\psi})$ , and standard deviation,  $\sigma(\hat{\psi})$ . Then it is easy to show that  $\hat{\psi}_{a;b}(\gamma)$  is a window function with center and standard deviation given by

$$m(\hat{\psi}_{a;b}) = \frac{m(\hat{\psi})}{a} \quad (127)$$

and

$$\sigma(\hat{\psi}_{a;b}) = \frac{\sigma(\hat{\psi})}{|a|}, \quad (128)$$

respectively. From Equations 126, 127 and 128 we have that in the frequency domain the analysis coefficient  $(\mathcal{W}_{\psi}s)(a, b)$  essentially depends on the values of  $\hat{s}(\gamma)$  for  $\gamma \in \left[ \frac{m(\hat{\psi})}{a} - \frac{\sigma(\hat{\psi})}{|a|}, \frac{m(\hat{\psi})}{a} + \frac{\sigma(\hat{\psi})}{|a|} \right]$ .

From Equations 122 through 128,  $(\mathcal{W}_{\psi}s)(a, b)$  yields a time-frequency representation of  $s(t)$ . The analysis coefficient  $(\mathcal{W}_{\psi}s)(a, b)$  depends on the time-frequency window

$$[am(\psi) + b - |a|\sigma(\psi), am(\psi) + b + |a|\sigma(\psi)] \times \left[ \frac{m(\hat{\psi})}{a} - \frac{\sigma(\hat{\psi})}{|a|}, \frac{m(\hat{\psi})}{a} + \frac{\sigma(\hat{\psi})}{|a|} \right].$$

In contrast to the short-time Fourier transform, both the time and frequency resolution in the wavelet transform vary in the time-frequency plane. The following properties derive from the time-frequency localization characteristics of the wavelet transform:

1. The ratio of the standard deviation,  $\sigma(\hat{\psi})/|a|$ , of the frequency window and its center frequency,  $m(\hat{\psi})/a$ , is given by  $\pm \frac{\sigma(\hat{\psi})}{m(\hat{\psi})}$ , which is independent of the location of the center frequency. This is called constant- $Q$  frequency analysis.



2. The time-frequency window narrows in time and widens in frequency for large center frequency  $m(\hat{\psi})/a$  (small  $a > 0$ ), and widens in time and narrows in frequency for small center frequency  $m(\hat{\psi})/a$  (large  $a > 0$ ).

Recall that the wavelet transform in Equation 119 can be viewed as the inner product of  $s(t)$  with the family of functions  $\psi_{a;b}(t)$  defined in Equation 121. Notice that this family of functions is obtained from a single basic wavelet by translating it by  $b$  and dilating it by  $a$ . The basic wavelet is usually referred to as the mother wavelet. The constant  $|a|^{-\frac{1}{2}}$  in Equation 121 is used for energy normalization purposes, that is  $\|\psi_{a;b}(t)\|_2^2 = \|\psi(t)\|_2^2$ . In general the mother wavelet  $\psi(t)$  is normalized to have its energy equal to one, that is  $\|\psi(t)\|_2^2 = 1$ . We assume this restriction from now on.

In practice the parameter  $a$  is restricted to  $a > 0$ , so that the center frequency of  $\hat{\psi}_{a;b}(\gamma)$  in Equation 127 is restricted to positive frequencies. In this context the parameter  $a$  is usually referred to as the scale parameter. In order to reconstruct  $s(t)$  from its wavelet transform restricted to  $a > 0$ , the Fourier transform of  $\psi(t)$  must satisfy a more strict admissibility condition than Equation 120 given by [74]

$$\int_0^{+\infty} \frac{|\hat{\psi}(\gamma)|^2}{\gamma} d\gamma = \int_0^{+\infty} \frac{|\hat{\psi}(-\gamma)|^2}{\gamma} d\gamma = \frac{1}{2} C_\psi < \infty. \quad (129)$$

For any  $\psi(t)$  satisfying Equation 129 we have the following reconstruction formula [74]

$$s(t) = \frac{2}{C_\psi} \int_{-\infty}^{+\infty} \int_0^{+\infty} (\mathcal{W}_\psi s)(a, b) \left\{ \frac{1}{\sqrt{a}} \psi \left( \frac{t-b}{a} \right) \right\} \frac{da}{a^2} db.$$

The wavelet transform is a redundant representation. Instead of computing  $(\mathcal{W}_\psi s)(a, b)$  for  $b \in \mathbb{R}$  and  $a \in \mathbb{R}^+$ , it is possible to choose an exponential sampling of the scale parameter,  $a = a_0^j$ ,  $j \in \mathbb{Z}$ ,  $a_0 \neq 0$ , such that the representation is complete and stable. Of particular interest is the sequence of scales where the elementary dilation step  $a_0 = 2$  because it leads to an octave by octave partitioning of the frequency domain.

In order to reconstruct  $s(t)$  from its wavelet transform restricted to  $b \in \mathbb{R}$  and  $a = a_0^j$ ,  $j \in \mathbb{Z}$ , the Fourier transform of  $\psi(t)$  must satisfy a more strict condition than Equation 129 given by [74]

$$A \leq \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(a_0^j \gamma) \right| \leq B, \quad (130)$$

for almost all  $\gamma \in \mathbb{R}$  for some constants  $A$  and  $B$  with  $0 < A \leq B < \infty$ . Reconstruction is possible by using the following formula [74]

$$s(t) = \sum_{j=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left\{ a_0^{-\frac{j}{2}} (\mathcal{W}_\psi s)(a_0^j, b) \right\} \left\{ a_0^{-j} \tilde{\psi} \left( \frac{t-b}{a_0^j} \right) \right\} db,$$

where the Fourier transform of  $\tilde{\psi}(t)$  is given by

$$\hat{\tilde{\psi}}(\gamma) = \frac{\hat{\psi}(\gamma)}{\sum_{j=-\infty}^{+\infty} |\hat{\psi}(a_0^j \gamma)|^2}.$$

If  $\psi(t)$  satisfies the so-called “stability condition” in Equation 130 with  $a_0 = 2$ , then  $\psi(t)$  is called a dyadic wavelet [74].

In addition to sampling the scale parameter  $a = a_0^j$ ,  $j \in \mathbb{Z}$ , it is possible to choose a sampling of the translation parameter  $b$  and still obtain a complete and stable representation. Intuitively, the discretization of  $b$  has to be chosen such that  $\psi_{a,b}(t)$  “covers” the entire time axis for each scale  $a = a_0^j$ ,  $j \in \mathbb{Z}$ . From Equation 124 we see that the standard deviation of  $\psi_{a,b}(t)$  is  $|a|$  times the standard deviation of  $\psi(t)$ . Hence, in order for  $\psi_{a,b}(t)$  to “cover” the entire time axis at scale  $a = a_0^j$ , the discretization of  $b$  has to be proportional to  $|a|$ . It follows that at scale  $a = a_0^j$  the translation has to be  $b = kb_0 a_0^j$ , where  $b_0 > 0$  is the elementary translation step. This leads to a discretized family of wavelets given by

$$\psi_{j,k}(t) = a_0^{-\frac{j}{2}} \psi \left( \frac{t - kb_0 a_0^j}{a_0^j} \right) = a_0^{-\frac{j}{2}} \psi(a_0^{-j} t - kb_0). \quad (131)$$

In order to reconstruct  $s(t)$  from its wavelet transform restricted to  $b = kb_0 a_0^j$ ,  $j, k \in \mathbb{Z}$  and  $a = a_0^j$ ,  $j \in \mathbb{Z}$ ,  $\psi(t)$  must satisfy a stricter condition than Equation 130. The admissibility condition for this reconstruction is the existence of  $A > 0$ ,  $B < \infty$  so that

$$A \|s\|_2^2 \leq \sum_{j,k \in \mathbb{Z}} |\langle s(t), \psi_{j,k}(t) \rangle|^2 \leq B \|s\|_2^2$$

for all  $s(t) \in L^2(\mathbb{R})$  [8]. In other words, the family of functions  $\psi_{j,k}(t)$  defined in Equation 131 constitute a frame. A brief review of frames is presented at the end of this section.

Given a frame of wavelets, reconstruction of  $s(t)$  from the  $\langle s(t), \psi_{j,k}(t) \rangle$  is possible by using the following formula

$$s(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \langle s(t), \psi_{j,k}(t) \rangle \tilde{\psi}^{j,k}(t), \quad (132)$$

where the family of functions  $\tilde{\psi}^{j,k}(t)$ ,  $j, k \in \mathbb{Z}$  is called the dual frame of  $\psi_{j,k}(t)$ ,  $j, k \in \mathbb{Z}$  (see the section below).

If the dual frame of  $\psi_{j,k}(t)$ ,  $j, k \in \mathbb{Z}$  is of the form

$$\tilde{\psi}^{j,k}(t) = \tilde{\psi}_{j,k}(t), \quad j, k \in \mathbb{Z},$$

where

$$\tilde{\psi}_{j,k}(t) = a_0^{-\frac{j}{2}} \tilde{\psi} \left( \frac{t - kb_0 a_0^j}{a_0^j} \right) = a_0^{-\frac{j}{2}} \tilde{\psi} (a_0^{-j} t - kb_0),$$

then Equation 132 is referred to as a wavelet series and  $\tilde{\psi}(t)$  is known as the dual wavelet of  $\psi(t)$ . Given a wavelet frame, the existence of a dual wavelet is important for computational purposes. Below, we overview the classification of wavelets according to orthogonality and highlight the existence of a dual wavelet in each case. A wavelet can be classified according to orthogonality as follows:

1. A wavelet  $\psi(t)$  is called semiorthogonal if it satisfies  $\langle \psi_{j,k}(t), \psi_{l,m}(t) \rangle = 0$ ,  $j \neq l, j, k, l, m \in \mathbb{Z}$ .

Remarks: It can be shown [74] that for every semiorthogonal wavelet,  $\psi(t)$ , there exist a dual wavelet,  $\tilde{\psi}(t)$ , such that the pair  $(\psi(t), \tilde{\psi}(t))$  satisfies the biorthogonality property

$$\langle \psi_{j,k}(t), \tilde{\psi}_{l,m}(t) \rangle = \delta_{j,l} \delta_{k,m}, j, k, l, m \in \mathbb{Z}. \quad (133)$$

2. A wavelet  $\psi(t)$  is called nonorthogonal if it is not a semiorthogonal wavelet.
3. A semiorthogonal wavelet  $\psi(t)$  is called orthogonal if it satisfies  $\langle \psi_{j,k}(t), \psi_{l,m}(t) \rangle = \delta_{j,l} \delta_{k,m}, j, k, l, m \in \mathbb{Z}$ .

Remarks: It can be shown that an orthogonal wavelet is self dual.

4. A nonorthogonal wavelet  $\psi(t)$  is called biorthogonal if there exists a dual wavelet,  $\tilde{\psi}(t)$ , such that the pair  $(\psi(t), \tilde{\psi}(t))$  satisfies the biorthogonality property in Equation 133.

The computation of Equation 132 can be a burden for nonorthogonal wavelets for which a dual wavelet does not exist. In such a case, it is advantageous to work with frames which are almost tight ("snug frames"), i.e., frames which have  $B/A - 1 \ll 1$ , because the  $\tilde{\psi}^{j,k}(t)$  can be approximated by  $\psi_{j,k}(t)$  (see the section below and [8] for more details).

## Frames

In this section we present a brief review of frames. For a more detailed and rigorous account of frames see [8].

A family of functions  $\varphi_j$ ,  $j \in \mathbb{J}$  in a Hilbert space  $\mathcal{H}$  is called a frame if there exist  $A > 0$ ,  $B < \infty$  so that for all  $f \in \mathcal{H}$ ,

$$A\|f\|^2 \leq \sum_{j \in \mathbb{J}} |\langle f, \varphi_j \rangle|^2 \leq B\|f\|^2, \quad (134)$$

where  $A$  and  $B$  are called the frame bounds. If the two frame bounds are equal, the frame is called a tight frame. In a tight frame we have, for all  $f \in \mathcal{H}$ ,

$$\sum_{j \in \mathbb{J}} |\langle f, \varphi_j \rangle|^2 = A\|f\|^2,$$

which implies

$$A \langle f, g \rangle = \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \langle \varphi_j, g \rangle,$$

or (at least in the weak sense [8]),

$$f = A^{-1} \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \varphi_j. \quad (135)$$

Although Equation 135 is reminiscent of the expansion of  $f$  into an orthonormal basis, it is important to note that frames, even tight frames, are not (orthonormal) bases. The family of functions  $\varphi_j$ ,  $j \in \mathbb{J}$  are typically not linearly independent. If the frame is tight, and if  $\|\varphi_j\| = 1$  for all  $j \in \mathbb{J}$ , then  $A = B$  gives the “redundancy ratio.” If this ratio equals to 1, then the tight frame is an orthonormal basis [8].

Equation 135 gives a trivial way to recover  $f$  from the  $\langle f, \varphi_j \rangle$ , if the frame is tight. Consider now recovering  $f$  from frames that are not tight. Let us define the frame operator  $F$  from  $\mathcal{H}$  to  $\ell^2(\mathbb{J})$  as

$$(Ff)_j = \langle f, \varphi_j \rangle.$$

Since  $\varphi_j$  constitute a frame, it follows from Equation 134 that  $\|Ff\|^2 \leq B\|f\|^2$ , that is,  $F$  is bounded, which means it is possible to find its adjoint operator  $F^*$  [8]. The adjoint operator  $F^*$  of  $F$  can be computed from the following relation

$$\begin{aligned} \langle F^*c, f \rangle &= \langle c, Ff \rangle = \sum_{j \in \mathbb{J}} c_j \overline{\langle f, \varphi_j \rangle} \\ &= \sum_{j \in \mathbb{J}} c_j \langle \varphi_j, f \rangle, \end{aligned}$$

so that

$$F^*c = \sum_{j \in \mathbb{J}} c_j \varphi_j,$$

at least in the weak sense [8]. From this it follows that

$$\sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \varphi_j = F^* F f.$$

Thus, the definition of  $F$  implies

$$\sum_{j \in \mathbb{J}} |\langle f, \varphi_j \rangle|^2 = \|F f\|^2 = \langle F^* F f, f \rangle.$$

Hence, in terms of  $F$ , the frame condition in Equation 134 can be written as

$$A I \leq F^* F \leq B I, \quad (136)$$

where  $I$  is the identity operator [8]. This implies, in particular, that  $F^* F$  is invertible (see Lemma 3.2.2 in [8]), and that the operator  $(F^* F)^{-1}$  satisfies,

$$B^{-1} I \leq (F^* F)^{-1} \leq A^{-1} I.$$

Applying the operator  $(F^* F)^{-1}$  to the family of functions  $\varphi_j$ ,  $j \in \mathbb{J}$ , leads to another family of functions denoted by  $\tilde{\varphi}_j$ ,  $j \in \mathbb{J}$  where

$$\tilde{\varphi}_j = (F^* F)^{-1} \varphi_j, \quad j \in \mathbb{J}.$$

The family of functions  $\tilde{\varphi}_j$ ,  $j \in \mathbb{J}$  also constitutes a frame with frame bounds  $B^{-1}$  and  $A^{-1}$  [8], that is

$$B^{-1} \|f\|^2 \leq \sum_{j \in \mathbb{J}} |\langle f, \tilde{\varphi}_j \rangle|^2 \leq A^{-1} \|f\|^2.$$

It can be verified (see Proposition 3.2.3 in [8]) that the associated frame operator  $\tilde{F}$  from  $\mathcal{H}$  to  $\ell^2(\mathbb{J})$ ,  $(\tilde{F} f)_j = \langle f, \tilde{\varphi}_j \rangle$  satisfies  $\tilde{F} = F(F^* F)^{-1}$ ,  $\tilde{F}^* \tilde{F} = (F^* F)^{-1}$ ,  $\tilde{F}^* F = I = F^* \tilde{F}$  and  $\tilde{F} F^* = F \tilde{F}^*$  is the orthogonal projection operator, in  $\ell^2(\mathbb{J})$ , onto  $\text{Ran}(F) = \text{Ran}(\tilde{F})$ . The family of functions  $\tilde{\varphi}_j$ ,  $j \in \mathbb{J}$  is called the dual frame of  $\varphi_j$ ,  $j \in \mathbb{J}$ . It is easy to verify that the dual frame of  $\tilde{\varphi}_j$ ,  $j \in \mathbb{J}$  is  $\varphi_j$ ,  $j \in \mathbb{J}$ . From  $\tilde{F}^* F = I = F^* \tilde{F}$  it follows that

$$\sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \tilde{\varphi}_j = f = \sum_{j \in \mathbb{J}} \langle f, \tilde{\varphi}_j \rangle \varphi_j. \quad (137)$$

Hence, we have a way to recover  $f$  from the  $\langle f, \varphi_j \rangle$ , where the only thing we need to do is to compute the  $\tilde{\varphi}_j = (F^* F)^{-1} \varphi_j$ . Note that, in general, if the frame is redundant, there exist other functions in  $\mathcal{H}$  that could equally well play the role of the  $\tilde{\varphi}_j$  and lead to a reconstruction formula. This follows from the fact that the  $\varphi_j$  are not linearly independent in the general case. Equation 137, however, yields the most “economical” representation of

$f$  in the following sense (see Proposition 3.2.4 in [8]): Consider the first half of Equation 137,  $f = \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \tilde{\varphi}_j$ . If  $f = \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle u_j$  for some other family of functions  $u_j$ ,  $j \in \mathbb{J}$ , then it can be shown [8] that  $\sum_{j \in \mathbb{J}} |\langle f, u_j \rangle|^2 \geq \sum_{j \in \mathbb{J}} |\langle f, \tilde{\varphi}_j \rangle|^2$ . Similarly, consider the second half of Equation 137,  $f = \sum_{j \in \mathbb{J}} \langle f, \tilde{\varphi}_j \rangle \varphi_j$ . If  $f = \sum_{j \in \mathbb{J}} c_j \varphi_j$  for some  $\mathbf{c} \in \ell^2(\mathbb{J})$ , and if not all  $c_j$  equal  $\langle f, \tilde{\varphi}_j \rangle$ , then it can be shown [8] that  $\sum_{j \in \mathbb{J}} |c_j|^2 \geq \sum_{j \in \mathbb{J}} |\langle f, \tilde{\varphi}_j \rangle|^2$ .

Computing the  $\tilde{\varphi}_j$  involves the inversion of  $F^*F$ . If  $B$  is close to  $A$  ( $r = B/A - 1 \ll 1$ ), then from Equation 136 we have that  $F^*F$  is “close” to  $\frac{A+B}{2} I$ . This implies that  $(F^*F)^{-1}$  is “close” to  $\frac{2}{A+B} I$ , and that  $\tilde{\varphi}_j$  is “close” to  $\frac{2}{A+B} \varphi_j$  [8]. This motivates the following reconstruction formula for  $f$  [8],

$$f = \frac{2}{A+B} \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \varphi_j + Rf, \quad (138)$$

where

$$R = I - \frac{2}{A+B} F^*F. \quad (139)$$

It follows from Equations 136 and 139 that  $-\frac{B-A}{B+A} I \leq R \leq \frac{B-A}{B+A} I$ . This implies that  $\|R\| \leq \frac{B-A}{B+A} = \frac{r}{2+r} < 1$ . If  $r$  is small, the rest term  $Rf$  in Equation 138 can be dropped, leading to a reconstruction formula for  $f$  which is accurate up to an  $L^2$ -error of  $\frac{r}{2+r} \|f\|$  [8]. Even if  $r$  is not small, it is possible to write an algorithm for the reconstruction of  $f$  with exponential convergence [8]. From Equation 139 we see that

$$F^*F = \frac{A+B}{2} (I - R).$$

This implies that

$$(F^*F)^{-1} = \frac{2}{A+B} (I - R)^{-1}.$$

Since  $\|R\| < 1$ , the series  $\sum_{k=0}^{\infty} R^k$  converges in norm, and its limit is  $(I - R)^{-1}$  [8]. It follows that

$$\tilde{\varphi}_j = (F^*F)^{-1} \varphi_j = \frac{2}{A+B} \sum_{k=0}^{\infty} R^k \varphi_j.$$

Note that the zeroth order term in the above equation leads exactly to Equation 138 with the rest term dropped. Better approximations can be obtained by truncating after  $N$  terms [8],

$$\tilde{\varphi}_j^N = \frac{2}{A+B} \sum_{k=0}^N R^k \varphi_j,$$

with

$$\left\| f - \sum_{j \in \mathbb{J}} \langle f, \varphi_j \rangle \tilde{\varphi}_j^N \right\| \leq \left( \frac{r}{2+r} \right)^{N+1} \|f\|,$$

which becomes exponentially small as  $N$  increases, since  $\frac{r}{2+r} < 1$ . In particular, the  $\tilde{\varphi}_j$  can be computed by an iterative algorithm [8],

$$\tilde{\varphi}_j^N = \frac{2}{A+B} \varphi_j + R \tilde{\varphi}_j^{N-1}.$$

Similarly,  $f$  can be computed by an iterative algorithm [8],

$$f = (F * F)^{-1} (F * F) f = \lim_{N \rightarrow \infty} f_N,$$

with

$$f_N = f_{N-1} + \frac{2}{A+B} \sum_{j \in \mathbb{J}} [\langle f, \varphi_j \rangle - \langle f_{N-1}, \varphi_j \rangle] \varphi_j.$$

### Wavelet Frames

Not all choices for  $\psi(t)$ ,  $a_0$ ,  $b_0$  lead to frames of wavelets, even if  $\psi(t)$  is admissible. In [8], Daubechies gives some general conditions on  $\psi(t)$ ,  $a_0$ ,  $b_0$  under which a frame is obtained and derives estimates for the frame bounds. These results can be summarized as follows (see Proposition 3.3.2 in [8]): If  $\psi(t)$ ,  $a_0$ , are such that

$$\begin{aligned} \inf_{1 \leq |\gamma| \leq a_0} \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(a_0^j \gamma) \right|^2 &> 0, \\ \sup_{1 \leq |\gamma| \leq a_0} \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(a_0^j \gamma) \right|^2 &< \infty, \end{aligned} \quad (140)$$

and if

$$\beta(\gamma_s) = \sup_{1 \leq |\gamma| \leq a_0} \sum_{j=-\infty}^{\infty} \left| \hat{\psi}(a_0^j \gamma) \right| \left| \hat{\psi}(a_0^j \gamma + \gamma_s) \right|$$

decays as fast as  $(1 + |\gamma_s|)^{-(1+\epsilon)}$ , with  $\epsilon > 0$ , then there exist  $b_{\text{thr}} > 0$  such that the  $\psi_{j,k}(t)$  constitute a frame for all choices  $b_0 < b_{\text{thr}}$ . For  $b_0 < b_{\text{thr}}$ , the following expressions are frame bounds for the  $\psi_{j,k}(t)$

$$A = \frac{1}{b_0} \left\{ \inf_{1 \leq |\gamma| \leq a_0} \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(a_0^j \gamma) \right|^2 - \sum_{\substack{k=-\infty \\ k \neq 0}}^{+\infty} \left[ \beta \left( \frac{2\pi}{b_0} k \right) \beta \left( -\frac{2\pi}{b_0} k \right) \right]^{1/2} \right\},$$

$$B = \frac{1}{b_0} \left\{ \sup_{1 \leq |\gamma| \leq a_0} \sum_{j=-\infty}^{+\infty} |\hat{\psi}(a_0^j \gamma)|^2 + \sum_{\substack{k=-\infty \\ k \neq 0}}^{+\infty} \left[ \beta \left( \frac{2\pi}{b_0} k \right) \beta \left( -\frac{2\pi}{b_0} k \right) \right]^{1/2} \right\}.$$

The condition on  $\beta(\gamma_s)$  and Equation 140 are satisfied if  $|\hat{\psi}(\gamma)| \leq C|\gamma|^\alpha(1 + |\gamma|)^{-\beta}$  with  $\alpha > 0$ ,  $\beta > \alpha + 1$ . It can also be verified that if the  $\psi_{j,k}(t)$  in Equation 131 constitute a frame with frame bounds  $A$  and  $B$  for some choice of  $a_0$ ,  $b_0$ , then the Fourier transform of  $\psi(t)$  satisfies

$$b_0 A \leq \sum_{j=-\infty}^{+\infty} |\hat{\psi}(a_0^j \gamma)|^2 \leq b_0 B \quad (141)$$

almost everywhere [74].

As mentioned before it is advantageous to set the elementary dilation step  $a_0 = 2$  because it leads to an octave by octave partitioning of the frequency domain. Hereafter we will assume this partitioning of the frequency domain unless stated otherwise.

### Multivoice Wavelet Frames

In Section 2.5.3 we mentioned that it is advantageous to use wavelet frames with  $B/A - 1 \ll 1$ , because the  $\tilde{\psi}^{j,k}(t)$  can be approximated by  $\psi_{j,k}(t)$ . This implies that the sum in Equation 141 should be almost constant for  $\gamma \neq 0$ . This imposes a strong restriction on  $\psi(t)$  that is not generally satisfied. In order to overcome this problem without having to give up too much freedom in choosing  $\psi(t)$  or its bandwidth, one can adopt the use of different voices per octave (i.e., suboctave sampling) [8]. This can be achieved by using different wavelets  $\psi_1(t), \psi_2(t), \dots, \psi_N(t)$  and looking at the frame  $\psi_{v;j,k}(t) = \frac{1}{\sqrt{2^j}} \psi_v \left( \frac{t}{2^j} - kb_0 \right)$ ,  $j, k \in \mathbb{Z}, v = 1, \dots, N$ . In [8], Daubechies gives the following expressions for the frame bounds of this multivoice frame

$$A = \frac{1}{b_0} \left\{ \inf_{1 \leq |\gamma| \leq 2} \sum_{v=1}^N \sum_{j=-\infty}^{+\infty} |\hat{\psi}_v(2^j \gamma)|^2 - \sum_{k \neq 0} \sum_{v=1}^N \left[ \beta_v \left( \frac{2\pi}{b_0} k \right) \beta_v \left( -\frac{2\pi}{b_0} k \right) \right]^{1/2} \right\},$$

$$B = \frac{1}{b_0} \left\{ \sup_{1 \leq |\gamma| \leq 2} \sum_{v=1}^N \sum_{j=-\infty}^{+\infty} |\hat{\psi}_v(2^j \gamma)|^2 + \sum_{k \neq 0} \sum_{v=1}^N \left[ \beta_v \left( \frac{2\pi}{b_0} k \right) \beta_v \left( -\frac{2\pi}{b_0} k \right) \right]^{1/2} \right\},$$

with

$$\beta_v(\gamma_s) = \sup_{1 \leq |\gamma| \leq 2} \sum_{j=-\infty}^{+\infty} |\hat{\psi}_v(2^j \gamma)| |\hat{\psi}_v(2^j \gamma + \gamma_s)|.$$



It is suggested in [8] that by choosing the  $\hat{\psi}_1(\gamma), \dots, \hat{\psi}_N(\gamma)$  to have slightly staggered frequency localization centers, coupled with good decay at  $\infty$ , one can achieve  $B/A - 1 \ll 1$ . One choice favored by several authors [78, 8, 79] is to take “fractionally” dilated versions of a single wavelet

$$\psi_v(t) = 2^{-(v-1)/N} \psi(2^{-(v-1)/N} t).$$

In this case  $\sum_{v=1}^N \sum_{j=-\infty}^{+\infty} |\hat{\psi}_v(2^j \gamma)|^2 = \sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^{j/N} \gamma)|^2$  which can be made almost constant by choosing  $N$  large enough [8].

#### 2.5.4 The À Trous and Mallat’s Algorithms

Recall from Section 2.5.3 that the wavelet transform of a signal  $s(t)$  is given by

$$(\mathcal{W}_\psi s)(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} s(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt, \quad s(t) \in L^2(\mathbb{R}), \quad a, b \in \mathbb{R}, \quad (a \neq 0).$$

The wavelet coefficients  $\langle s(t), \psi_{j,k}(t) \rangle$  in Equation 132 for  $a_0 = 2$ ,  $b_0 = 1$  can be obtained by sampling the wavelet transform of  $s(t)$  with  $a = 2^j$ ,  $j \in \mathbb{Z}$  and  $b = 2^j k$ ,  $j, k \in \mathbb{Z}$ . We will refer to these coefficients as the decimated wavelet “series” coefficients and will denote them by

$$\begin{aligned} (C_\psi s)_{j,k} &= (\mathcal{W}_\psi s)(2^j, 2^j k) \\ &= \langle s(t), \psi_{j,k}(t) \rangle, \end{aligned} \tag{142}$$

where  $\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - k\right)$  is the generator of a family of decimated wavelets. We will refer to the operation of computing the above wavelet coefficients as the decimated wavelet transform.

For many applications the above sampling is inadequate. In particular, a finer grid is needed when computing the wavelet transform for applications that require “good” time–frequency localization. This can be achieved by computing an undecimated wavelet transform and the addition of voices [66]. The undecimated wavelet transform of a signal  $s(t)$  is obtained by sampling the wavelet transform of  $s(t)$  with  $a = 2^j$ ,  $j \in \mathbb{Z}$  and  $b = k$ ,  $k \in \mathbb{Z}$ . We will refer to these coefficients as the undecimated wavelet “series” coefficients and will denote them by

$$\begin{aligned} (R_\psi s)_{j,k} &= (\mathcal{W}_\psi s)(2^j, k) \\ &= \langle s(t), \psi_k^j(t) \rangle, \end{aligned} \tag{143}$$

where  $\psi_k^j(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-k}{2^j}\right)$  is the generator of a family of undecimated wavelets.

In what follows we will review two separately motivated implementations of the wavelet transform, the *algorithme à trous* [13] and Mallat's [41] multiresolution decomposition. These algorithms are both special cases of a single filter bank structure, the discrete wavelet transform [66], the behavior of which is governed by one's choice of filters. The à trous algorithm was originally devised as a computationally efficient implementation of the wavelet transform, however it is more properly viewed as a nonorthogonal multiresolution decomposition for which the discrete wavelet transform computes a "sampled" continuous wavelet transform (wavelet "series" coefficients) exactly [66]. On the other hand, Mallat's algorithm was devised as an orthogonal multiresolution decomposition [41] that when initialized properly computes a "sampled" continuous wavelet transform [66]. For a more detailed and rigorous account of these algorithms see [41, 13, 66].

The following additional terminology and notation will be used. The discretized wavelet transform (with sampling period  $T = 1$ ) is given by

$$(w_\psi s)(a, b) = \frac{1}{\sqrt{a}} \sum_k s(k) \overline{\psi\left(\frac{k-b}{a}\right)}. \quad (144)$$

The discretized decimated wavelet "series" coefficients are given by

$$(c_\psi s)_{j,k} = (w_\psi s)(2^j, 2^j k) = \frac{1}{\sqrt{2^j}} \sum_n s(n) \overline{\psi\left(\frac{n}{2^j} - k\right)}. \quad (145)$$

Similarly, the discretized undecimated wavelet "series" coefficients are given by

$$(r_\psi s)_{j,k} = (w_\psi s)(2^j, k) = \frac{1}{\sqrt{2^j}} \sum_n s(n) \overline{\psi\left(\frac{n-k}{2^j}\right)}. \quad (146)$$

Signals and filters in bold face will be treated as vectors. The  $k$ th element of vector  $\mathbf{f}$  is denoted by  $[\mathbf{f}]_k = f_k$ . The symbol  $\dagger$  will be used for the mirror filter  $[\mathbf{f}^\dagger]_k = f_k^\dagger = \overline{f_{-k}}$ . The decimator operator is denoted by a matrix

$$\begin{aligned} D_{k,m} &= \delta(2k - m) \\ &= \delta_{2k,m} \end{aligned}$$

where  $\delta_{k,m}$  is the Kronecker delta and  $\delta(k) = \delta_{k,0}$ . The dilation operator is denoted by a matrix

$$\begin{aligned} U_{k,m} &= \delta(k - 2m) \\ &= \delta_{k,2m}. \end{aligned}$$

Convolution of  $\mathbf{f}$  and  $\mathbf{s}$  is denoted by  $[\mathbf{f} * \mathbf{s}]_k = \sum_m f_{k-m} s_m$ . Convolution followed by decimation becomes  $[\mathbf{D}(\mathbf{f} * \mathbf{s})]_k = \sum_m D_{k,m} [\mathbf{f} * \mathbf{s}]_m = [\mathbf{f} * \mathbf{s}]_{2k} = \sum_m f_{2k-m} s_m$ . The

following shorthand will be used for the convolution of  $\mathbf{f}$  and  $\mathbf{s}$ ,  $[\mathbf{f} * \mathbf{s}] = \mathbf{F}\mathbf{s}$ , where  $F_{m,n} = f_{m-n}$ . It is easy to verify that  $[\mathbf{D}(\mathbf{f} * \mathbf{s})] = \mathbf{D}\mathbf{F}\mathbf{s}$ . The symbol  $\dagger$  will also be used for the Hermitian transpose of matrices  $[\mathbf{F}^\dagger]_{m,n} = \overline{f_{n-m}} = f_{m-n}^\dagger$ . For discrete signals the translation operator is denoted as

$$[T_m \mathbf{s}]_n = s_{m-n},$$

and for continuous signals as

$$(T_\tau s)(t) = s(t - \tau).$$

The  $z$  transform of a discrete signal  $s_n$  is given by

$$\hat{s}(z) = \sum_n s_n z^{-n}.$$

Finally, we state the following two identities [66], which will be used later. For any  $\mathbf{F}$  of the form  $F_{m,n} = f_{m-n}$  we have that

$$[(\mathbf{D}\mathbf{F})^j]_{n,k} = [(\mathbf{D}\mathbf{F})^j]_{0,k-2^j n} \quad (147)$$

and

$$\sum_k [(\mathbf{D}\mathbf{F})^j]_{n,k} z^k = z^{2^j n} \prod_{r=0}^{j-1} \hat{f}(z^{2^r}). \quad (148)$$

### The Decimated À Trouis Algorithm

The à trous algorithm was originally devised as an efficient implementation for computing wavelet series coefficients. To achieve this, the wavelet series coefficients in Equations 142 and 143 are approximated by their discretized counterparts in Equations 145 and 146, respectively. As a starting point consider implementing Equation 145. Clearly, as  $j$  increases  $\psi(t)$  must be sampled at progressively more points, creating a large computational burden [13, 66]. The solution proposed by [13] is to approximate nonintegral points via a finite filter  $\mathbf{f}^\dagger$ . As an example, let  $\mathbf{f}^\dagger$  be the filter  $\frac{1}{\sqrt{2}}(0.5, 1.0, 0.5)$ . Then,

$$\begin{aligned} \sqrt{2} \sum_k f_{n-2k}^\dagger \psi(k) &= \begin{cases} \psi(\frac{n}{2}), & \text{for } n \text{ even,} \\ \frac{1}{2}(\psi(\frac{n-1}{2}) + \psi(\frac{n+1}{2})), & \text{for } n \text{ odd,} \end{cases} \\ &\approx \psi\left(\frac{n}{2}\right), \end{aligned}$$

approximates a sampling of  $\psi(t/2)$ . Let  $\mathbf{g}$  be a filter defined by  $g_n = \overline{\psi(-n)}$ . Then the above interpolation can be computed by first dilating  $\mathbf{g}^\dagger$  and then convolving the result

with a filter  $\mathbf{f}^\dagger$  which leaves the even points fixed and interpolates to get the odd points [66]. The condition that  $\mathbf{f}$  be the identity on even points is sufficiently important that merits the following definition: A lowpass filter  $\mathbf{f}$  is said to be an à trous filter if it satisfies

$$f_{2k} = \delta(k)/\sqrt{2}. \quad (149)$$

The  $\sqrt{2}$  is simply a convenient means of including the normalization factor of Equation 145 in the filter.

It can be verified [66] that the interpolation operation can be written as follows

$$\begin{aligned} [\mathbf{f}^\dagger * (\mathbf{U}\mathbf{g}^\dagger)]_n &= [\mathbf{F}^\dagger \mathbf{U}\mathbf{g}^\dagger]_n \\ &= \sum_k f_{n-2k}^\dagger g_k^\dagger \\ &= \sum_k f_{n-2k}^\dagger \psi(k) \\ &\approx \frac{1}{\sqrt{2}} \psi\left(\frac{n}{2}\right). \end{aligned}$$

This result and Equation 145 leads to the following approximation for the discretized decimated wavelet series coefficients at  $j = 1$ ,

$$\begin{aligned} (c_\psi s)_{1,k} &= \frac{1}{\sqrt{2}} \sum_n s(n) \overline{\psi\left(\frac{n}{2} - k\right)} \\ &\approx \sum_n s_n \sum_m \overline{f_{n-2k-2m}^\dagger} \bar{g}_m^\dagger \\ &= \sum_m g_{k-m} \sum_n f_{2m-n} s_n \\ &= [\mathbf{g} * \mathbf{D}(\mathbf{f} * \mathbf{s})]_k \\ &= [\mathbf{G}\mathbf{D}\mathbf{F}\mathbf{s}]_k, \end{aligned}$$

where  $[\mathbf{s}]_n = s_n = s(n)$ . Proceeding inductively, one can find the following approximation for the discretized decimated wavelet series coefficients for all  $j$  [66]

$$(c_\psi s)_{j,k} \approx [\mathbf{G}(\mathbf{D}\mathbf{F})^j \mathbf{s}]_k. \quad (150)$$

Remarkably, the right-hand side in the above equation is the discrete decimated wavelet transform of  $\mathbf{s}$  at level  $j$  [66], which we denote as follows<sup>5</sup>

$$[\mathbf{c}_j(\mathbf{s})]_k = [\mathbf{G}(\mathbf{D}\mathbf{F})^j \mathbf{s}]_k. \quad (151)$$

---

<sup>5</sup>The notation  $[\mathbf{c}_j]_k = c_{j,k}$  will be used as a shorthand for  $[\mathbf{c}_j(\mathbf{s})]_k$ .

Equations 150 and 151 yield the following formula for the discretized decimated wavelet series coefficients

$$(\mathbf{c}_\psi s)_{j,k} \approx [\mathbf{c}_j(\mathbf{s})]_k. \quad (152)$$

The discrete decimated wavelet transform can be computed iteratively over  $j$  as follows

$$\begin{aligned} [\mathbf{s}_{j+1}]_k &= \sum_n f_{2k-n} [\mathbf{s}_j]_n \\ &= [\mathbf{D}(\mathbf{f} * \mathbf{s}_j)]_k \\ \\ [\mathbf{c}_j]_k &= \sum_n g_{k-n} [\mathbf{s}_j]_n \\ &= [\mathbf{g} * \mathbf{s}_j]_k \end{aligned} \quad (153)$$

where  $[\mathbf{s}_0]_k = s_k$ . Except for decimation of the output the à trous algorithm described in [13] is given by Equations 152 and 153.

The original à trous algorithm made no statements regarding the accuracy of the approximation in Equation 150 or even the discretization from Equation 119 to Equation 144. A major step forward towards treating this question lies in the results of [66] outlined in Section 2.5.5.

### Mallat's Algorithm

In this section we review Mallat's multiresolution algorithm. For a more detailed and rigorous account of this algorithm see [41, 80, 66]. Mallat's algorithm has basically the same structure as Equations 153, that is,

$$\begin{aligned} [\mathbf{s}_{j+1}]_k &= [\mathbf{D}(\mathbf{h} * \mathbf{s}_j)]_k \\ \\ [\mathbf{d}_{j+1}]_k &= [\mathbf{D}(\mathbf{g} * \mathbf{s}_j)]_k \end{aligned} \quad (154)$$

where  $[\mathbf{s}_0]_k = s_k$ . In keeping with the literature, the lowpass filter is denoted by  $\mathbf{h}$  instead of  $\mathbf{f}$ . The constraints on  $\mathbf{h}$  and  $\mathbf{g}$  which ensure an orthonormal multiresolution analysis are [41, 66]:

#### 1. Perfect reconstruction

$$\sum_j h_{2j-m} \bar{h}_{2j-n} + g_{2j-m} \bar{g}_{2j-n} = \delta_{m,n}, \quad (155)$$

2. Orthogonality of  $\mathbf{h}$  and  $\mathbf{g}$  with respect to even shifts

$$\sum_j h_{2m-j} \bar{g}_{2n-j} = 0, \quad (156)$$

3. Bandpass condition on  $\mathbf{g}$

$$\sum_n g_n = 0, \quad (157)$$

4. Lowpass condition on  $\mathbf{h}$

$$\sum_n h_n = \sqrt{2}. \quad (158)$$

In matrix notation Equations 155 and 156 may be written as follows

$$(\mathbf{H}^\dagger \mathbf{U})(\mathbf{D}\mathbf{H}) + (\mathbf{G}^\dagger \mathbf{U})(\mathbf{D}\mathbf{G}) = \mathbf{I}, \quad (159)$$

$$(\mathbf{D}\mathbf{H})(\mathbf{G}^\dagger \mathbf{U}) = \mathbf{0}, \quad (160)$$

where  $H_{m,n} = h_{m-n}$ . Multiplying Equation 159 on the left by  $\mathbf{D}\mathbf{H}$  and using Equation 160 we see that

$$(\mathbf{D}\mathbf{H})(\mathbf{H}^\dagger \mathbf{U}) = \mathbf{I}.$$

It follows that  $\mathbf{h}$  and its shifted versions by even shifts form an orthonormal set

$$\sum_j h_{2m-j} \bar{h}_{2n-j} = \delta_{m,n}, \quad m, n \in \mathbb{Z}.$$

Similarly, multiplying Equation 159 on the left by  $\mathbf{D}\mathbf{G}$  and using Equation 160 we have that

$$(\mathbf{D}\mathbf{G})(\mathbf{G}^\dagger \mathbf{U}) = \mathbf{I}.$$

It follows that  $\mathbf{g}$  and its shifted versions by even shifts also form an orthonormal set

$$\sum_j g_{2m-j} \bar{g}_{2n-j} = \delta_{m,n}, \quad m, n \in \mathbb{Z}.$$

These two results together with Equation 156 imply that Equations 154 is an orthogonal decomposition of the discrete signal  $\mathbf{s}_j$ . Furthermore, Equation 155 implies that

$$[\mathbf{s}_j]_k = [(\mathbf{H}^\dagger \mathbf{U})\mathbf{s}_{j+1} + (\mathbf{G}^\dagger \mathbf{U})\mathbf{d}_{j+1}]_k.$$

From Equations 155–158 it follows that Equation 154 represents a wavelet decomposition as described below [41, 66].

Define a scaling function  $\phi(t)$  with Fourier transform given by

$$\hat{\phi}(\omega) = \prod_{r=1}^{\infty} \frac{1}{\sqrt{2}} \overline{\hat{h}\left(e^{j\frac{\omega}{2^r}}\right)}.$$

It follows that

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}} \overline{\hat{h}\left(e^{j\frac{\omega}{2}}\right)} \hat{\phi}\left(\frac{\omega}{2}\right)$$

which in the time domain takes the form

$$\phi(t) = \sum_k \bar{h}_{-k} \sqrt{2} \phi(2t - k).$$

Therefore, the dilates and translates of  $\phi(t)$ ,

$$\phi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t}{2^j} - n\right)$$

have the property [66]

$$\phi_{j+1,k}(t) = \sum_n \bar{h}_{2k-n} \phi_{j,n}(t).$$

Define a wavelet function by

$$\psi(t) = \sum_k \bar{g}_{-k} \sqrt{2} \phi(2t - k). \quad (161)$$

Then, using Equation 156 and the above properties, one can show that the family of wavelets

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - n\right)$$

are orthonormal, and that the  $\mathbf{d}_j$  are the coefficients of the expansion of  $s(t)$  in terms of  $\psi_{j,k}(t)$ , that is

$$[\mathbf{d}_j]_k = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} s(t) \overline{\psi\left(\frac{t}{2^j} - k\right)} dt,$$

provided that

$$[\mathbf{s}_0]_k = \int_{-\infty}^{+\infty} s(t) \overline{\phi(t - k)} dt.$$

Therefore, Mallat's algorithm when initialized properly (see the equation above) computes exactly a “sampled” continuous wavelet transform [66].

## The Undecimated À Trouis Algorithm

As mentioned earlier for many applications the time-frequency sampling grid provided by the decimated wavelet transform is inadequate when “good” time-frequency localization is required. One way to overcome this limitation is by computing an undecimated wavelet transform. Following the same approach described above, the wavelet series coefficients in Equation 143 can be approximated by their discretized counterpart in Equation 146. From Equations 145 and 146 it follows that  $(c_\psi s)_{j,k} = (w_\psi s)(2^j, 2^j k)$  and  $(r_\psi s)_{j,k} = (w_\psi s)(2^j, k)$  should coincide at  $k = 0$ . Utilizing this fact, one can obtain the  $k$ th undecimated wavelet series coefficient  $(r_\psi s)_{j,k}$  by translating the signal back by  $k$  and computing the discretized decimated wavelet series coefficient  $(c_\psi(T_{-k}s))_{j,0}$ . This result together with Equation 152 yields the following approximation for computing the discretized undecimated wavelet series coefficients

$$\begin{aligned} (r_\psi s)_{j,k} &= (c_\psi(T_{-k}s))_{j,0} \\ &\approx [c_j(T_{-k}\mathbf{s})]_0, \end{aligned} \quad (162)$$

where  $[\mathbf{s}]_n = s_n = s(n)$ . Remarkably, the right-hand side in the above equation is the discrete undecimated wavelet transform of  $\mathbf{s}$  at level  $j$  [66], which we denote as follows<sup>6</sup>

$$[r_j(\mathbf{s})]_k = [c_j(T_{-k}\mathbf{s})]_0. \quad (163)$$

Equations 162 and 163 yield the following formula for the discretized undecimated wavelet series coefficients

$$(r_\psi s)_{j,k} \approx [r_j(\mathbf{s})]_k. \quad (164)$$

A few observations about the decimated and undecimated discrete wavelet transforms are in order. Notice that in general  $c_j$  is not translation invariant

$$\begin{aligned} [c_j(T_m \mathbf{s})]_k &= \sum_n [G(D\mathbf{F})^j]_{k,n} [\mathbf{s}]_{n-m} \\ &= \sum_n [G(D\mathbf{F})^j]_{k,n+m} [\mathbf{s}]_n \\ &\neq \sum_n [G(D\mathbf{F})^j]_{k-m,n} [\mathbf{s}]_n. \end{aligned}$$

However, if one replaces  $m$  by  $2^j m$  in the above equation and uses Equation 147, the last step becomes an equality [66]

$$[c_j(T_{2^j m} \mathbf{s})]_k = [c_j(\mathbf{s})]_{k-m}. \quad (165)$$

---

<sup>6</sup>The notation  $[r_j]_k = r_{j,k}$  will be used as a shorthand for  $[r_j(\mathbf{s})]_k$



Thus, translating  $\mathbf{s}$  by  $2^j m$ , translates octave  $j$  by  $m$ .

On the other hand, it can be verified that  $\mathbf{r}_j$  is translation invariant

$$\begin{aligned} [\mathbf{r}_j(T_m \mathbf{s})]_k &= [\mathbf{c}_j(T_{m-k} \mathbf{s})]_0 \\ &= [\mathbf{r}_j(\mathbf{s})]_{k-m}. \end{aligned}$$

Also, from Equations 163 and 165 it follows that sampling  $r_{j,k}$  every  $2^j$  points produces exactly  $c_{j,k}$ , that is,

$$c_{j,k} = r_{j,2^j k}.$$

An iterative algorithm for computing the discrete undecimated wavelet transform can be obtained by taking  $z$  transforms [66]. From Equations 151 and 163

$$\begin{aligned} \sum_k r_{j,k} z^{-k} &= \sum_m \sum_k [\mathbf{G}(\mathbf{D}\mathbf{F})^j]_{0,m} s_{m+k} z^{-k} \\ &= \sum_m [\mathbf{G}(\mathbf{D}\mathbf{F})^j]_{0,m} z^m \hat{s}(z). \end{aligned}$$

Applying Equation 148 to the above equation it follows that

$$\sum_k r_{j,k} z^{-k} = \hat{g}(z^{2^j}) \prod_{r=0}^{j-1} \hat{f}(z^{2^r}) \hat{s}(z)$$

where  $j = 0$  is understood to mean there are no factors of  $\hat{f}$ .

It is easy to see that  $\mathbf{U}^j \mathbf{f}$  is  $\mathbf{f}$  with  $2^j - 1$  zeros inserted between every pair of filter coefficients and that its  $z$  transform is  $\hat{f}(z^{2^j})$ . Therefore the discrete undecimated wavelet transform can be computed iteratively over  $j$  as follows

$$[\mathbf{s}_{j+1}]_k = [(\mathbf{U}^j \mathbf{f}) * \mathbf{s}_j]_k$$

$$[\mathbf{r}_j]_k = [(\mathbf{U}^j \mathbf{g}) * \mathbf{s}_j]_k$$

where  $[\mathbf{s}_0]_k = s_k$ . The above result together with Equation 164 is essentially the à trous algorithm described in [13].

### 2.5.5 The À Trous Algorithm as an Exact Wavelet Transform

In the previous section we analyzed two separately motivated implementations of the wavelet transform. We learned that the à trous algorithm was originally devised as a computationally efficient approximation of the “sampled” wavelet transform  $(C_\psi s)_{j,k}$  whose

implementation can be viewed as a discrete wavelet transform for which the filter  $\mathbf{f}$  is an interpolator satisfying Equation 149 and the filter  $\mathbf{g}$  is obtained by sampling the wavelet function  $\psi(t)$ . We also learned that Mallat's algorithm is a discrete wavelet transform with filters  $\mathbf{h}$  and  $\mathbf{g}$  subject to the constraints in Equations 155–158 that when initialized properly computes exactly a “sampled” continuous wavelet transform for the wavelet function  $\psi(t)$  in Equation 161. Next we will review the results of [66] which show that the à trous algorithm can be viewed as a nonorthogonal multiresolution decomposition for which the discrete wavelet transform computes a “sampled” continuous wavelet transform exactly.

Define a scaling function  $\phi(t)$  with Fourier transform given by

$$\hat{\phi}(\omega) = \prod_{r=1}^{\infty} \frac{1}{\sqrt{2}} \overline{\hat{f}\left(e^{j\frac{\omega}{2^r}}\right)}, \quad (166)$$

or equivalently

$$\phi(t) = \lim_{j \rightarrow \infty} \sum_k [(\overline{\mathbf{D}\mathbf{F}})^j]_{0,k} \sqrt{2^j} \chi(2^j - k),$$

where

$$\chi(t) = \begin{cases} 1 & \text{for } t \in [-1/2, 1/2), \\ 0 & \text{otherwise.} \end{cases}$$

For normalization purposes  $\hat{f}(e^{j\omega})|_{\omega=0} = \sqrt{2}$ , which implies

$$\sum_k f_k = \sqrt{2} \quad (167)$$

and

$$\int_{-\infty}^{+\infty} \phi(t) dt = \hat{\phi}(0) = 1.$$

Suitable regularity conditions for the inverse Fourier transform of the product in Equation 166 to converge to a reasonable behaved function may be found in [81, 82, 66]. One can verify [66] that the dilates and translates of  $\phi(t)$  have the property

$$\begin{aligned} \phi_{j+1,k}(t) &= \sum_n [\overline{\mathbf{D}\mathbf{F}}]_{k,n} \phi_{j,n}(t). \\ &= \sum_n \bar{f}_{2k-n} \phi_{j,n}(t). \end{aligned} \quad (168)$$

Furthermore, if  $\mathbf{f}$  is a finite filter, then  $\phi(t)$  has finite support [66].

Define the wavelet function

$$\psi'(t) = \sum_k \bar{g}_k \phi(t+k) = \sum_k g_k^\dagger \phi(t-k). \quad (169)$$

The recursion in Equation 168 implies that

$$\phi_{j,k}(t) = \sum_n [(\overline{D\mathbf{F}})^j]_{k,n} \phi_{0,n}(t).$$

This expression along with Equation 169 yields

$$\psi'_{j,k}(t) = \sum_n [\overline{G(D\mathbf{F})^j}]_{k,n} \phi_{0,n}(t).$$

Using the above expression in Equation 145 the discretized decimated wavelet “series” coefficients take the form

$$\begin{aligned} (c_{\psi' s})_{j,k} &= \sum_l s(l) \overline{\psi'_{j,k}(l)} \\ &= \sum_l s(l) \sum_n [\overline{G(D\mathbf{F})^j}]_{k,n} \phi_{0,n}(l) \\ &= \sum_n [G(D\mathbf{F})^j]_{k,n} \sum_l s(l) \overline{\phi_{0,n}(l)} \\ &= \sum_n [G(D\mathbf{F})^j]_{k,n} \sum_l s(l) \overline{\phi(l-n)}. \end{aligned}$$

Define

$$[s]_n = \sum_l s(l) \overline{\phi(l-n)}, \quad (170)$$

then the above equation becomes

$$(c_{\psi' s})_{j,k} = [G(D\mathbf{F})^j s]_k,$$

which is the discrete decimated wavelet transform of the discrete signal  $s_n$  defined in Equation 170. Therefore, using the initialization in Equation 170 the discrete decimated wavelet transform computes exactly the discretized decimated wavelet “series” coefficients  $(c_{\psi' s})_{j,k}$ . Similarly, using the above expression for  $\psi'_{j,k}(t)$  in Equation 142 we see that the decimated wavelet “series” coefficients take the form

$$\begin{aligned} (C_{\psi' s})_{j,k} &= \int_{-\infty}^{+\infty} s(t) \overline{\psi'_{j,k}(t)} dt \\ &= \int_{-\infty}^{+\infty} s(t) \left( \sum_n [G(D\mathbf{F})^j]_{k,n} \overline{\phi_{0,n}(t)} \right) dt \\ &= \sum_n [G(D\mathbf{F})^j]_{k,n} \int_{-\infty}^{+\infty} s(t) \overline{\phi_{0,n}(t)} dt \\ &= \sum_n [G(D\mathbf{F})^j]_{k,n} \int_{-\infty}^{+\infty} s(t) \overline{\phi(t-n)} dt. \end{aligned}$$

Using

$$[\mathbf{s}]_n = \int_{-\infty}^{+\infty} s(t) \overline{\phi(t-n)} dt \quad (171)$$

the above equation becomes

$$(C_{\psi'} s)_{j,k} = [\mathbf{G}(\mathbf{D}\mathbf{F})^j \mathbf{s}]_k,$$

which is the discrete decimated wavelet transform of the discrete signal  $s_n$  defined in Equation 171. Therefore, using the initialization in Equation 171 the discrete decimated wavelet transform computes exactly the decimated wavelet “series” coefficients  $(C_{\psi'} s)_{j,k}$ . Notice that 170 is the discretized version of Equation 171.

Finally, we review the significance of the à trous condition in Equation 149. One may show that a filter  $\mathbf{f}$  is an á trous filter *iff*  $\phi(n) = \delta(n)$  [66]. Using this observation it can be shown that if  $\mathbf{f}$  is à trous then (1) Equation 170 becomes  $[\mathbf{s}]_n = s(n)$  and the discrete decimated wavelet transform of the sampled signal  $s(n)$  yields the exact discretized decimated wavelet “series” coefficients  $(c_{\psi'} s)_{j,k}$  and (2) the wavelet  $\psi'(t)$  defined in Equation 169 satisfies  $\psi'(n) = g_n^\dagger$  [66].

In summary, given  $\mathbf{f}$  satisfying Equations 149 and 167,  $\mathbf{g}$  defined by  $g_n^\dagger = \psi(n)$  (where  $\psi(t)$  is an arbitrary wavelet), and given the initialization in Equation 171, the discrete wavelet transform computes an exact “sampled” continuous wavelet transform of  $s(t)$  using the wavelet function  $\psi'(t)$  defined in Equation 169. Furthermore, if there is sufficient regularity,  $\psi'(t)$  and  $\psi(t)$  will be “close” since they coincide on the integers up to the length of  $\mathbf{g}$ .

### 2.5.6 The Sine–Gabor Wavelet

It is well known that the modulated Gaussian or Gabor function is the only function that achieves optimum time-frequency localization. The question arises as to whether the Gabor function,

$$\psi_G(t) = K_G e^{-t^2/2\sigma_0^2} e^{-j\omega_0 t},$$

can be used in the context of wavelet analysis and synthesis. Unfortunately, this is not possible because the Gabor function is not an admissible wavelet (for details on the admissibility condition for wavelets see Section 2.5.3.) Therefore, optimum time-frequency localization via the wavelet transform is not possible. However, a modified version of the Gabor function, known as the Morlet wavelet [8],

$$\psi_M(t) = K_M e^{-t^2/2\sigma_0^2} \left( e^{-j\omega_0 t} - e^{-\sigma_0^2 \omega_0^2/2} \right),$$

satisfies the wavelet admissibility condition, and for  $\sigma_0 = 1$ ,  $\omega_0 = \pi (2/\ln 2)^{1/2}$ , achieves nearly optimum time-frequency localization and generates a frame of wavelets (for more details about wavelet frames see Section 2.5.3.) In fact, the error incurred in approximating the Morlet wavelet with a Gabor function under these conditions is negligible [8]. Although Gabor functions are not wavelets in the strict sense, the term “Gabor wavelet” is used throughout the literature. We will refer to Gabor functions as Gabor wavelets provided that  $e^{-\sigma_0^2 \omega_0^2/2}$  is negligible.

Having a frame of wavelets,  $\psi_{j,k}(t)$ ,  $j, k \in \mathbb{Z}$ , is a desirable property because it allows one to completely characterize and reconstruct a signal  $s(t)$  from its discrete wavelet coefficients  $\langle s(t), \psi_{j,k}(t) \rangle$  (see Section 2.5.3.) In practice it is desirable to have a frame of wavelets for which both analysis and synthesis can be computed efficiently.

We show that the sine-Gabor function (complex part of the Morlet wavelet) achieves nearly optimum time-frequency localization and at the same time generates a frame of wavelets. Furthermore, we show that both analysis and synthesis using the sine-Gabor wavelet can be computed efficiently.

The fact that the linear combination of two wavelets is also a wavelet gives us an easy way to show that the sine-Gabor function is a wavelet. By taking the real and imaginary parts of the Morlet wavelet we see that

$$\operatorname{Re} \psi_M(t) = K_M e^{-x^2/2\sigma_0^2} \left( \cos(\omega_0 t) - e^{-\sigma_0^2 \omega_0^2/2} \right),$$

and

$$\operatorname{Im} \psi_M(t) = -K_M e^{-x^2/2\sigma_0^2} \sin(\omega_0 t).$$

It follows that the sine-Gabor function is indeed a wavelet. It also follows that the cosine-Gabor function is not a wavelet for otherwise the Gabor function would be a wavelet.

The question as to whether the sine-Gabor wavelet generates a frame of wavelets and the efficient computation of both analysis and synthesis using the sine-Gabor wavelet are addressed below.

### 2.5.7 The Sine-Gabor Wavelet Frame

We showed in Section 2.5.6 that sine-Gabor functions are wavelets. In this section we study the time-frequency localization characteristics of the sine-Gabor wavelets and also establish that under certain conditions they generate wavelet frames.

The sine-Gabor wavelet is given by

$$\psi(t) = K_\psi e^{-x^2/2\sigma_0^2} \sin(\omega_0 t), \tag{172}$$

where

$$K_\psi = \pi^{-1/4} \left( \frac{2}{\sigma_0(1 - e^{-\sigma_0^2 \omega_0^2})} \right)^{1/2}$$

was chosen so that  $\|\psi\|_2 = 1$ . The time and frequency resolution of the sine-Gabor wavelet are given by the root mean square extent of the function and its Fourier transform, respectively (see Equation 111 in Section 2.5.2.) It follows from Equations 172 and 111 that the time-resolution of the sine-Gabor wavelet is given by

$$\sigma(\psi) = \left( \frac{\sigma_0^2(1 - e^{-\sigma_0^2 \omega_0^2}(1 - 2\sigma_0^2 \omega_0^2))}{2(1 - e^{-\sigma_0^2 \omega_0^2})} \right)^{1/2}.$$

The Fourier transform of the sine-Gabor wavelet is given by

$$\hat{\psi}(\omega) = K_\psi \frac{e^{-\sigma_0^2(\omega - \omega_0)^2/2} - e^{-\sigma_0^2(\omega + \omega_0)^2/2}}{2j},$$

where

$$K_{\hat{\psi}} = K_\psi \sqrt{2\pi\sigma_0}.$$

Computing the root mean square extent of  $\hat{\psi}(\omega)$  as defined in Equation 111 would lead to misleading results. It can be verified that if a function  $\psi(t)$  satisfies the wavelet admissibility condition and is also a window function then its Fourier transform  $\hat{\psi}(\omega)$  must vanish at the origin (see Equation 125 in Section 2.5.3.) In addition if the wavelet  $\psi(t)$  is real we have that  $\hat{\psi}(\omega) = \overline{\hat{\psi}(-\omega)}$ , so that  $|\hat{\psi}(\omega)|$  is a symmetric function around the origin. It is therefore more suitable to restrict our attention to positive frequencies for this class of wavelets and use

$$\sigma(\hat{\psi}) = \left\{ \frac{2}{\|\hat{\psi}\|_2^2} \int_0^\infty (\omega - m(\hat{\psi}))^2 |\hat{\psi}(\omega)|^2 d\omega \right\}^{1/2},$$

for the frequency-resolution, where

$$m(\hat{\psi}) = \frac{2}{\|\hat{\psi}\|_2^2} \int_0^\infty \omega |\hat{\psi}(\omega)|^2 d\omega.$$

It is easy to verify that the sine-Gabor wavelet satisfies all the conditions mentioned above and that its frequency-resolution is given by

$$\sigma(\hat{\psi}) = \left( \frac{1 + 2\sigma_0^2 \omega_0^2 - e^{-\sigma_0^2 \omega_0^2}}{2\sigma_0^2(1 - e^{-\sigma_0^2 \omega_0^2})} - (m(\hat{\psi}))^2 \right)^{1/2}$$

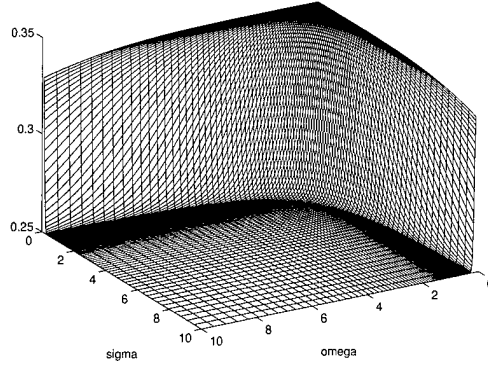


Figure 35: Square of the product of the time and frequency resolution for the sine-Gabor wavelet as a function of  $\sigma_0$  and  $\omega_0$ .

where

$$m(\hat{\psi}) = \frac{\omega_0 \operatorname{erf}(\sigma_0 \omega_0)}{1 - e^{-\sigma_0^2 \omega_0^2}}.$$

Figure 35 shows the square of the product of the time and frequency resolution for the sine-Gabor wavelet as a function of  $\sigma_0$  and  $\omega_0$ . For  $\sigma_0 \omega_0 \geq 5/2$ , the sine-Gabor wavelet is within 1% of the optimum time-frequency localization lower bound given by the uncertainty principle,  $\sigma^2(\psi) \sigma^2(\hat{\psi}) \geq 1/4$  (for more details see Equation 118 in Section 2.5.2.)

Next, we establish that the sine-Gabor function does indeed generate a frame of wavelets. General conditions on a wavelet  $\psi(t)$  under which a frame is obtained and estimates for the frame bounds are discussed in Section 2.5.3. A wavelet  $\psi(t)$  generates a frame of wavelets if the following condition is satisfied

$$|\hat{\psi}(\omega)| \leq C|\omega|^\alpha (1 + |\omega|)^{-\beta}, \quad (173)$$

with  $\alpha > 0$ ,  $\beta > \alpha + 1$ .

The magnitude of the Fourier transform of the sine-Gabor wavelet may be written as follows

$$\begin{aligned} |\hat{\psi}(\omega)| &= \frac{K_{\hat{\psi}}}{2} \left(1 - e^{-2\sigma_0^2 \omega_0 |\omega|}\right) e^{-\sigma_0^2 (|\omega| - \omega_0)^2 / 2} \\ &\leq \frac{K_{\hat{\psi}}}{2} (2\sigma_0^2 \omega_0 |\omega|)^{1/2} e^{-\sigma_0^2 (|\omega| - \omega_0)^2 / 2} \\ &\leq \frac{K_{\hat{\psi}}}{2} \frac{(2\sigma_0^2 \omega_0 |\omega|)^{1/2}}{1 + \sigma_0^2 (|\omega| - \omega_0)^2 / 2}, \end{aligned}$$

where the first and second inequalities are established by using

$1 - e^{-\omega} \leq \omega^\gamma$ ,  $\forall \omega > 0$ ,  $0 \leq \gamma \leq 1$ , and  $e^{-\omega^2/2} \leq \frac{1}{1+\omega^2/2}$ ,  $\forall \omega$ , respectively. It follows that the sine-Gabor wavelet satisfies Equation 173 with  $\alpha = 0.4$  and  $\beta = 1.85$  for some  $C$  sufficiently large.

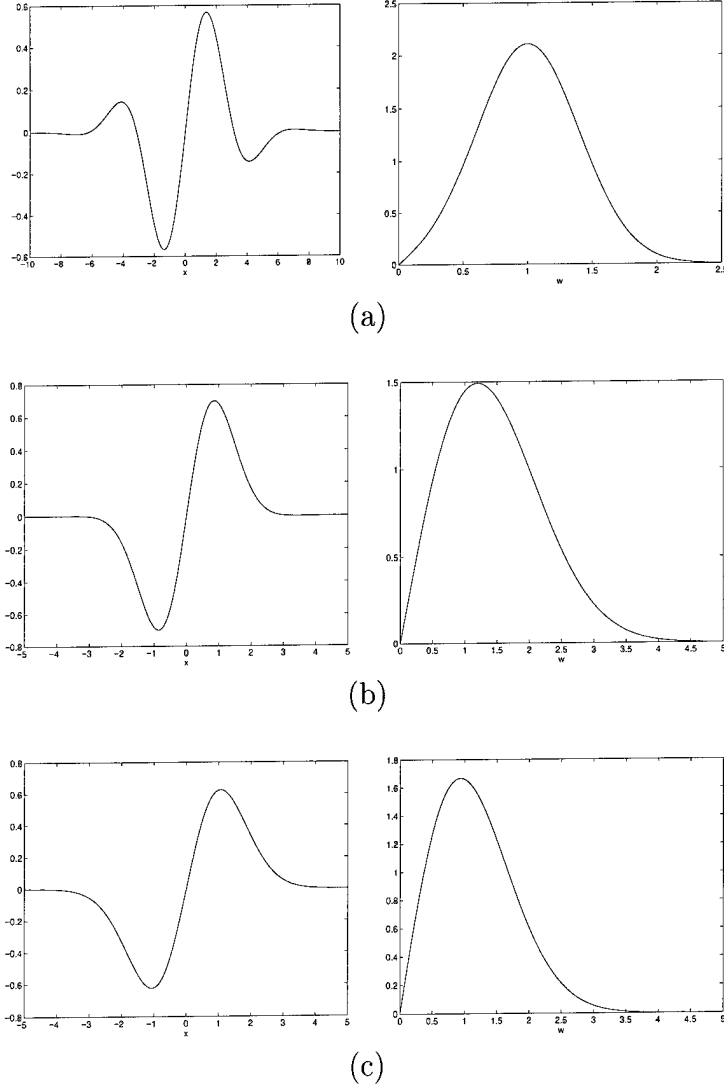


Figure 36: Sine-Gabor wavelet (left) and magnitude of its Fourier transform (right) with (a)  $\sigma_0 = 5/2$ ,  $\omega_0 = 1$ , (b)  $\sigma_0 = 1$ ,  $\omega_0 = 1$  and (c)  $\sigma_0 = 1.0657$ ,  $\omega_0 = 0.0299$ .

### Frame Bounds

Table 6 and Figure 36 show the estimated frame bounds and sine-Gabor wavelet for various values of  $\sigma_0$  and  $\omega_0$ . The results in Table 6(a) show that it is possible to obtain a frame of wavelets by using the sine-Gabor wavelet and at the same time achieve nearly optimum time-frequency localization,  $\sigma^2(\psi)\sigma^2(\hat{\psi}) = 0.2525$ . Table 6(b) shows that by trading off time-frequency localization,  $\sigma^2(\psi)\sigma^2(\hat{\psi}) = 0.3297$ , the frame bounds can be made tighter. As mentioned in Section 2.5.3, snug frame bounds are a desirable property because the dual wavelet can be approximated by  $\frac{2}{A+B}\psi(t)$ . The error between the original signal and its reconstruction using the above approximation is given by  $\frac{r}{2+r}\|f\|_2$ , where  $r = B/A - 1$  (see Section 2.5.3 for more details.) For the sine-Gabor wavelet with  $\sigma_0 = 1$ ,



Table 6(a). Frame bounds  
with  $\sigma_0 = 5/2$ ,  $\omega_0 = 1$

$b_0$	$A$	$B$	$B/A$
1.00	4.3873	5.6329	1.2839
1.25	3.5099	4.5063	1.2839
1.50	2.9216	3.7586	1.2865
1.75	2.4089	3.3169	1.3769
2.00	1.6016	3.4085	2.1281
2.25	0.3963	4.0572	10.2385

Table 6(b). Frame bounds  
with  $\sigma_0 = 1$ ,  $\omega_0 = 1$

$b_0$	$A$	$B$	$B/A$
0.50	8.5076	8.6988	1.0225
0.75	5.6715	5.7995	1.0226
1.00	4.1918	4.4114	1.0524
1.25	2.8464	4.0362	1.4180
1.50	1.2635	4.4720	3.5395

Table 6(c). Frame bounds  
with  $\sigma_0 = 1.0657$ ,  $\omega_0 = 0.0299$

$b_0$	$A$	$B$	$B/A$
0.75	7.2030	7.3285	1.0174
1.00	5.3999	5.4988	1.0183
1.25	4.2456	4.4733	1.0536
1.50	3.1604	4.1054	1.2990
1.75	1.9203	4.3075	2.2432
2.00	0.6524	4.7969	7.3523

Table 6: Frame bounds for the sine-Gabor wavelet with (a)  $\sigma_0 = 5/2$ ,  $\omega_0 = 1$ , (b)  $\sigma_0 = 1$ ,  $\omega_0 = 1$ , and (c)  $\sigma_0 = 1.0657$ ,  $\omega_0 = 0.0299$ . The elementary dilation parameter  $a_0 = 2$  in all cases.

Table 7. Frame bounds  
with  $\sigma_0 = 1.0657$ .

$b_0$	$A$	$B$	$B/A$
0.75	7.2043	7.3298	1.0174
1.00	5.4008	5.4997	1.0183
1.25	4.2465	4.4739	1.0536
1.50	3.1615	4.1055	1.2986
1.75	1.9218	4.3071	2.2412
2.00	0.6541	4.7962	7.3328

Table 7: Frame bounds for the first derivative of a Gaussian with  $\sigma_0 = 1.0657$ .

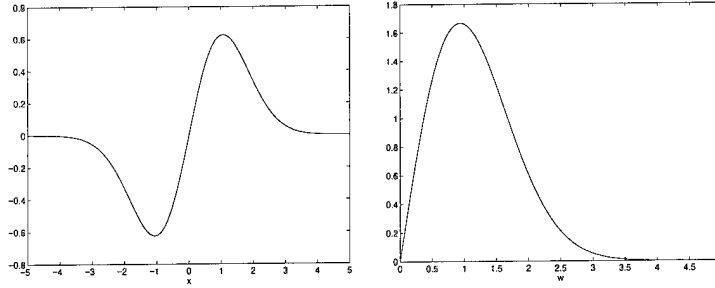


Figure 37: First derivative of a Gaussian (left) and the magnitude of its Fourier transform (right) with  $\sigma_0 = 1.0657$ .

$\omega_0 = 1$ , the error is 2.55% for  $a_0 = 2, b_0 = 1$ . The parameters for the sine-Gabor wavelet used in Table 6(c) were obtained by using a Simplex search method to find the minimum of the difference between  $B$  and  $A$  as a function of  $\sigma_0$  and  $\omega_0$ . The error in this case is 0.90% for  $a_0 = 2, b_0 = 1$ . This reduced error is at the expense of time-frequency localization,  $\sigma^2(\psi)\sigma^2(\hat{\psi}) = 0.3401$ . An important observation follows from this last example: For  $\sigma_0 = 1.0657, \omega_0 = 0.0299$  the sine-Gabor wavelet is nearly equal<sup>7</sup> (up to a sign) to the first derivative of a Gaussian with  $\sigma_0 = 1.0657$ . Therefore, the first derivative of a Gaussian (Canny's approximation to the optimal step edge detector [19]) can be used to completely characterize and reconstruct a signal from its discrete wavelet coefficients. Figure 37 shows the first derivative of a Gaussian and the magnitude of its Fourier transform. Table 7 shows the frame bounds for the first derivative of a Gaussian with  $\sigma_0 = 1.0657$ . The advantage of the sine-Gabor function over the first derivative of a Gaussian is that by changing the product of  $\sigma_0\omega_0$ , one is able to trade off between Canny's criteria for designing optimal step edge detectors [19], that is, the product of signal-to-noise ratio and localization,  $\Sigma\Delta$ , and the multiple response constraint,  $r$ . This is a property of Canny's optimal step edge detector that is not possible to achieve with the first derivative of a Gaussian. Table 8 shows the frame bounds for the sine-Gabor wavelet for  $0.75 \leq \sigma_0\omega_0 \leq 3.75$ .

### 2.5.8 Summary

In this chapter we have presented the tools that will allow us to compute a multivoice undecimated wavelet transform exactly for an arbitrary wavelet such as the sine-Gabor wavelet. Furthermore, we have shown that the sine-Gabor wavelet is indeed a wavelet frame. However, we found that having nearly optimum time-frequency localization and obtaining a tight frame at the same time is not possible. Both constraints are important because they allow for "good" time-frequency localization and efficient reconstruction,

<sup>7</sup>It is assumed that both functions have the same  $L^2$ -norm.

Table 8. Frame bounds for  
 $0.75 \leq \sigma_0 \leq 3.75$ ,  $\omega_0 = 1$ .

$\sigma_0$	$A$	$B$	$B/A$
3.75	2.5933	6.8524	2.6423
3.50	2.9659	6.5120	2.1956
3.25	3.3385	6.2107	1.8603
3.00	3.7031	5.9576	1.6088
2.75	4.0546	5.7629	1.4213
2.50	4.3873	5.6329	1.2839
2.25	4.6920	5.5627	1.1856
2.00	4.9418	5.5248	1.1180
1.75	5.0866	5.4618	1.0738
1.50	5.0580	5.2931	1.0465
1.25	4.7875	4.9433	1.0325
1.00	4.1918	4.4114	1.0524
0.75	2.8851	4.0837	1.4154

Table 8: Frame bounds for the sine-Gabor wavelet for  $0.75 \leq \sigma_0 \omega_0 \leq 3.75$ . The elementary dilation parameter  $a_0 = 2$  and the elementary translation step  $b_0 = 1$  in all cases.

respectively. We have also learned that a better (tighter) frame can be obtained by adding voices. This will not only ease the reconstruction but is also a requirement for applications in time–frequency analysis. Our current work is concerned with extending the sine-Gabor wavelet to the multivoice framework and with the computation of multivoice undecimated wavelet transforms. Our future work includes wavelet reconstruction in the multivoice/undecimated framework as well as extending our results to the analysis of two dimensional data including digital mammograms.

## 2.6 Circular Mass Recognition Based on the Hough Transform

### 2.6.1 Introduction

The Hough Transform (HT) is a standard method for shape recognition in digital images [83, 84]. It was first applied to the recognition of straight lines [85, 86] and later extended to circles [87], ellipses [88] and arbitrarily shaped objects [89]. Its advantages include robustness to noise, robustness to shape distortions and to occlusions/missing parts of an object. Its main disadvantage is the fact that computational and storage requirements of the algorithm increase as a power of the dimensionality of the curve. This means that for straight lines the computational complexity and storage requirements are  $O(n^2)$ , for circles  $O(n^3)$  and for ellipses  $O(n^5)$ .

We begin with a review of the problem of circular mass recognition using the HT.

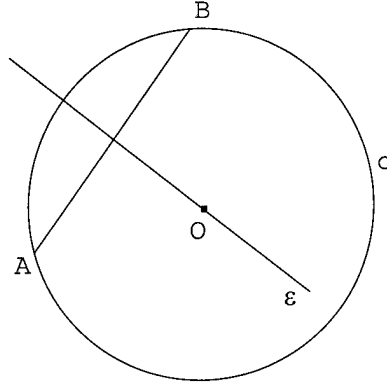


Figure 38: Demonstration of the fact that the center of circle  $c$ ,  $O$ , belongs to the line that perpendicularly bisects the segment defined by points  $A$  and  $B$  (both points belong to the circumference of the circle).

Even though there have been attempts towards the recognition of circles via the standard 3D HT [90], it has been recognized that there is a need for a decomposition of the search space, to simplify the problem both in terms of computation and storage. Previous algorithms have been based on the property that the normal to a point of the circumference of a circle passes through the center of the circle [91]. This approach works well for high signal to noise ratios and/or simple environments. As the signal to noise ratio decreases, the accuracy of the gradient estimation decreases [92, 93]. The fact that gradient-based methods are heavily dependent upon the accuracy of the gradient estimation explains why they are not robust to noise. Another disadvantage is that, sometimes, the edge detector of choice does not provide gradient information. A comparative study of various HT based techniques for circle recognition has been performed by Yuen *et al.* [94].

The approach taken in our study was to decompose the 3D search into a 2D HT and 1D radius histogramming. For the first part, instead of relying on a 2D gradient-based HT we developed a 2D bisection based HT. The property we exploit is that the line that perpendicularly bisects any chord of a circle passes through its center (see Figure 38). Experimentation with synthetic data demonstrated that our approach is more robust to noise than gradient-based techniques. The price we pay for robustness is an increase in the computational requirements due to the labeling of the connected components that our method requires.

The second part of the algorithm, 1D radius histogramming, is used to validate the existence of identified circles and calculate their radius. We show that extracting information from the radius histogram is not a trivial task and we devise a filtering technique that solves this problem.

Next, we discuss the steps of our algorithm. In this section we give details about the

implementation of the algorithm and analyze the effects of various factors, such as digitization of the image, discretization of the parameter space, noise, etc., on the accuracy and the computational efficiency of the method. In Section 2.6.3, we provide sample demonstrations of the algorithm with synthetic and real images. In the first part of this section, we use synthetic data to demonstrate the robustness of the technique and compare it with the gradient-based 2D HT. In the second part of this subsection, we present results of processing a digitized phantom and real mammograms from the Mammographic Image Analysis Society (MIAS): MiniMammography Database. Finally, in Section 2.6.5, we present some conclusions and summary.

## 2.6.2 Algorithm

### A 2D Hough Transform based on bisection

After edge detection, the resulting connected components are labeled. Each connected component is expressed as a chain of the coordinates of its points. We connect pairs of points, of the same component, using a sliding window (see Figure 39). If the coordinates of the points are  $A(x_A, y_A)$  and  $B(x_B, y_B)$ , the equation of the line that perpendicularly bisects  $AB$  is

$$y = \frac{x_A - x_B}{y_A - y_B}x + \frac{x_A^2 + y_A^2 - x_B^2 - y_B^2}{2(y_A - y_B)}. \quad (174)$$

All points (members of the parameter space) belonging to this line have their votes increased by one. Highly voted points provide an indication of the existence of circles. These points are the centers of such circles.

Since we are dealing with digital circles the points of their circumference are affected by digitization and, therefore, do not exactly satisfy the standard circle equation:

$$r^2 = (x - x_0)^2 + (y - y_0)^2, \quad (175)$$

where  $r$  is the radius of the circle and  $(x_0, y_0)$  are the coordinates of the center of the circle. Furthermore, there is a need for a discretization of the parameter space for two reasons:

1. *Computational efficiency.* It is impossible to account for all the circles that may exist in the image, and
2. *The line that perpendicularly bisects a chord of the circle is highly unlikely to pass through its center.* Due to the digitization of the image, pixels belonging to a digital circle do not exactly satisfy Equation 175 and, therefore, their chords do not coincide with the chords of a true circle.

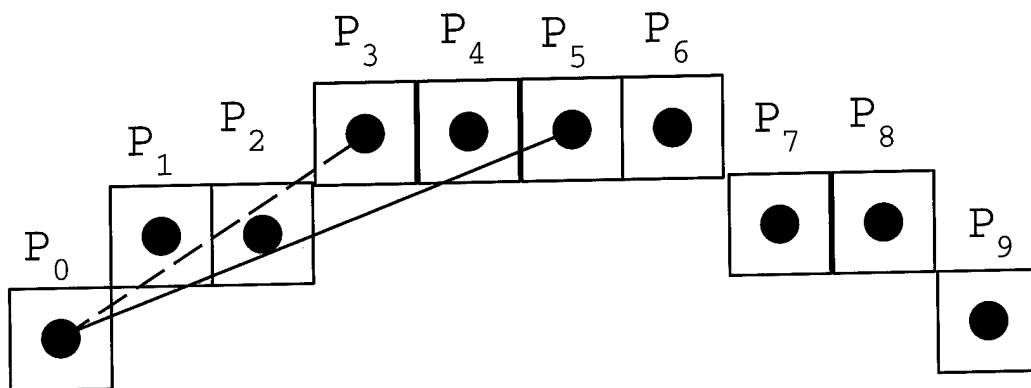


Figure 39: Graphic illustrating our implementation of a 2D HT for circle center localization. Points  $P_0$  and  $P_3$  of the connected component  $P_0, P_1, \dots, P_9$ , are assumed to belong to the same circle and all points that belong to the line that perpendicularly bisects the line segment they define receive votes. The same process is repeated for the pairs  $P_1-P_4, P_2-P_5$  etc. For this particular case the length of the window is equal to three. If one chooses to connect points  $P_0-P_5, P_1-P_6$ , etc., the length of the window is equal to five.

The discretization of the parameter space makes the voting scheme robust to noise and errors induced during the edge detection and other preprocessing operations. On the other hand, it introduces some error in the determination of the position of centers of circles. The larger the size of the members of the parameter space, the higher the uncertainty of the estimation of the position of the center of the circle. In our implementation the parameter space is congruent with the *image space* (this is dictated by the adopted parameterization) and the size of a member of the parameter space is the same as the size of a pixel. This choice is reasonable but not necessary. For example, one may choose a more robust scheme where the size of a member of the parameter space is  $2h$  by  $2h$  ( $h$  is the size of the pixel). Such a discretization will not only increase the uncertainty of the estimation of the position of the center of the circle but will also increase the probability of getting accidental peaks. Its advantage is that there is a smaller chance of missing real centers.

Simple detection of centers of circles is not enough. The reason is that:

1. Highly voted pixels provide only an indication of the existence of a circle. There is a need for verification of this hypothesis. Highly voted local maxima may be formed accidentally, and
2. The need for a determination of the third parameter of the circle (its radius  $r$ ).

Next we describe an efficient method towards the verification of the existence of a circle, and the extraction of its radius via radius-histogramming.

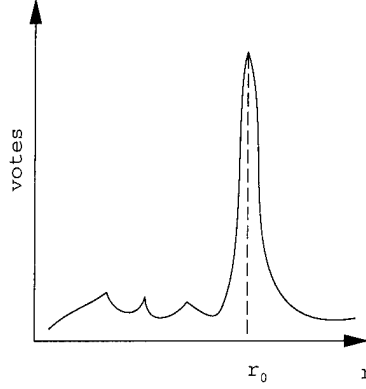


Figure 40: The radius histogram for the center of a circle. All bins get a small number of votes except for the bin to which the circle belongs. A sharp maximum appears for the bin that contains the circle.

### Analysis of radius histogram

After detecting possible centers, we can use the histogram of the distances of all feature points from the centers to verify the existence of circles and extract their radii (see Figure 40). In the ideal case (continuous circle, exactly determined center) the analysis of the radius histogram would be an easy goal. This is the case because circles would show up as sharp local maxima in a noisy background in the radius histogram. The reasons why this does not happen in a discrete space are:

1. Digitization/discretization errors. Even for the case where we can exactly determine the center of the circle, the digitization of the circle along with the digitization of the image combined with the discretization of the histogram make almost certain that the votes of the pixels will be spread to a number of neighboring bins of the histogram. As one can see from Figure 41, this effect is similar to the spreading of the straight line standard HT [95]. Under the boundary quantization scheme [96], it can be shown that *digital* circles can be bounded by two *Euclidean* circles with the same center and radii that differ by  $h$  [97]. This is illustrated in Figure 41 where a digital circle is bounded by continuous circles  $c_1$  and  $c_2$ . If  $\Delta r$  is the size of the bins of the histogram the maximum spreading for a circle is equal to:

$$n_r = \left\lfloor \frac{h}{\Delta r} \right\rfloor + 2, \quad (176)$$

where the symbol  $\lfloor z \rfloor$  denotes the largest integer less than  $z$ ,

2. Distortion of the shape due to imperfections in the image formation (i.e. breast compression), errors during the edge detection stage, and imperfections in the boundary of the mass.

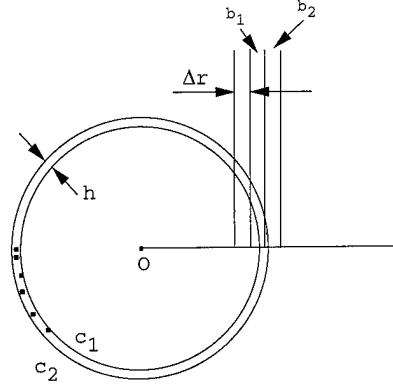


Figure 41: All pixels belonging to a digital circle can be bounded by two concentric circles,  $c_1$  and  $c_2$ , whose radii differ by  $h$ . For the digital circle of this figure, all pixels are spread between two neighboring bins,  $b_1$  and  $b_2$ .

3. Noisy pixels and other objects (tissues) that appear in the neighborhood of the circle.
4. Pixels missing from the boundary. This can happen because of occlusions, missing parts of the object, etc.
5. Errors in the localization of the center during the previous steps, mainly during the 2D Hough voting step. If the localization of the center of the circle is not accurate, two peaks will appear in the radius histogram of the estimated center (Figure 42). This has been discussed by Yuen *et al.* [94]. For a digital image if the error in the estimation is small, say less than two pixels, there appears a single extended peak. The length of the peak is 3 to 4 bins. If the error in the position estimation of the center is larger, two local maxima appear in the histogram. As the error increases so does the distance between the two peaks. Figure 42 illustrates these ideas.
6. The number of pixels belonging to a digital circle and a digital ring increases *almost* linearly. Kulpa [97] showed that as  $r \rightarrow +\infty$

$$P_c(r) = 4\sqrt{2}r, \quad (177)$$

where  $P_c$  is the number of points of a digital circle with radius  $r$ . He also provided experimental evidence that good approximation of the number of pixels belonging to a digital ring (the digital object bounded by two continuous circles whose radii differ by  $h$ ) is given by

$$P_r(r) = 2\pi r. \quad (178)$$

We should emphasize here, that the equivalent of a bin in a digital image is a digital ring and *not* a digital circle.



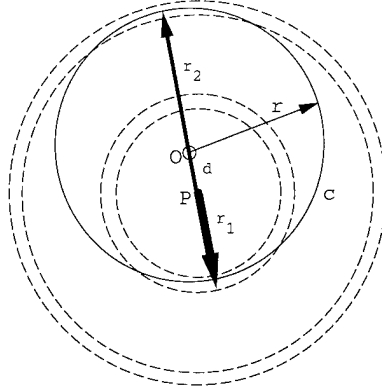


Figure 42: If an inaccurate estimate of the center of the circle is provided by the first step, the radius histogram will give two peaks. In this example, the estimated center is point  $P$ , while the true center is point  $O$ . Obviously, these peaks will appear at bins with distance  $r_1$  and  $r_2$  from the estimated center.

All these factors complicate the assessment of the radius histogram. Our approach towards the extraction of information from the radius histogram is based on the previously presented arguments.

The filter we propose is given by the following equation

$$t = \frac{1}{4\sqrt{2}r} \left[ -\frac{3r}{2(r-2)} \quad 1 \quad 1 \quad 1 \quad -\frac{3r}{2(r+2)} \right], \quad (179)$$

The details of the derivation of the coefficients of this filter are presented in Appendix A.3. The divisor on the right hand side of Equation 179 is equal to the number of pixels belonging to a circle of radius  $r$  (see Equation 177). We should add that the proposed filtering technique gives an unbiased estimator of the *normalized, unoccluded part* of the circumference of the circle. In other words, if no circular features exist in the image and the image is corrupted with uniform noise, the members of the histogram will have zero mean.

The resulting filtered histogram is searched for local peaks whose values exceed a certain threshold. Due to the normalization of the filtered histogram the threshold does not depend on the radius of the circle.

### 2.6.3 Results

In this section we present sample results from experiments applying the algorithm with real and synthetic digital images. Our purpose was to study the variation of robustness for the bisection-based HT with increasing noise and compare it with gradient-based HTs. As a measure of robustness, we used the robustness ratio  $P$ , defined by

$$P = \frac{v_{peak}}{v_{total}}, \quad (180)$$

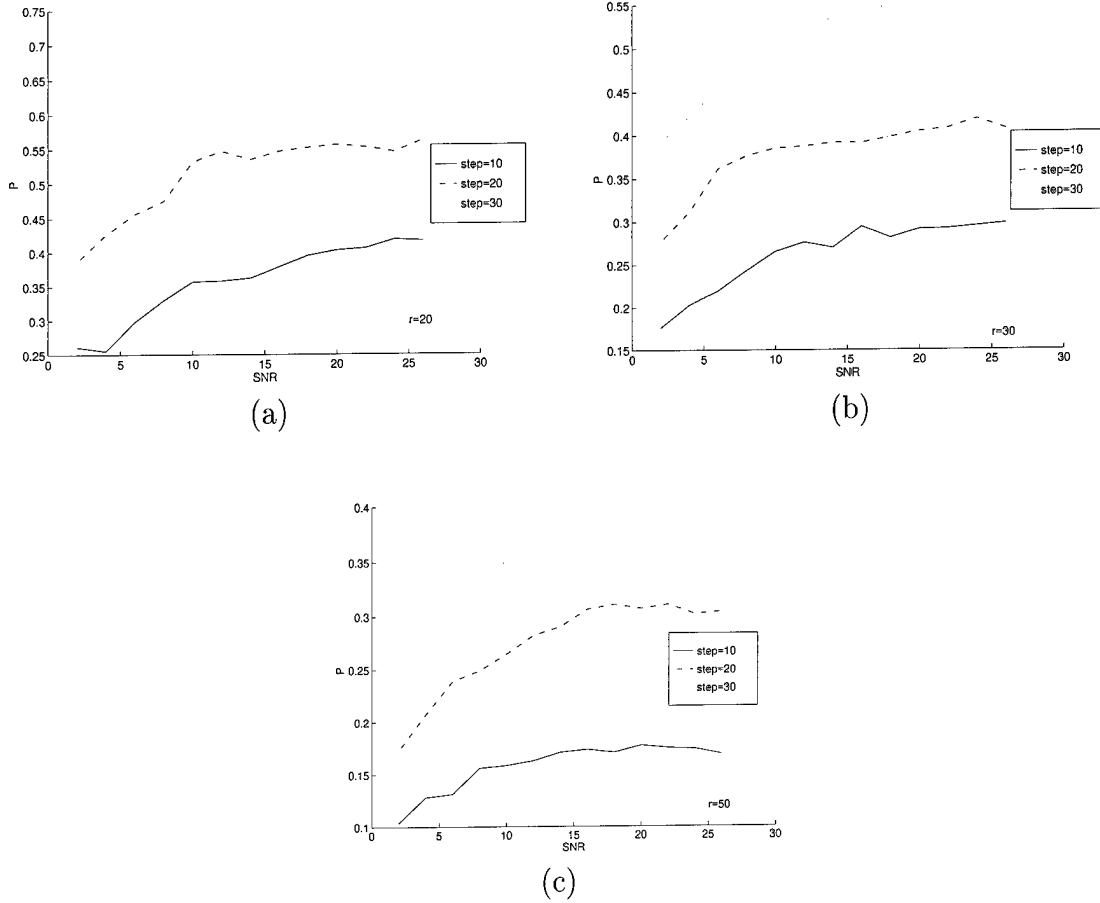


Figure 43: Comparison of the robustness  $P$  as a function of the  $SNR$  for three circles of radii: (a)  $r = 20$ ; (b)  $r = 30$ ; and (c)  $r = 50$  pixels. As the step becomes larger  $P$  increases.

where  $v_{peak}$  was the number of votes of the peak cell and  $v_{total}$  was the number of pixels of the circumference of the circle. Obviously, due to the digitization of the image, even in the ideal case,  $P$  is less than 1.

### Length of window

The first thing studied is the dependence of  $P$  with respect to the size of the window slid along the boundary of an object. Figure 39 shows examples of windows of different sizes slid along a digital curve. Figure 43 shows plots of  $P$  for three circles of radii 20, 30, and 50 pixels. We observed that as the radius increases so does  $P$ . We also noticed that as the size of the window increases,  $P$  also increased. Any theoretical analysis that would indicate the optimum window size is difficult due to the fact that errors in the coordinates of the two extreme points of the window are not independent of each other. Such a study has been presented by Amir [98]. The assumptions this study relied upon, were independent errors in the coordinates of the two extreme points of the window and an  $a$

*priori* knowledge of the radius of the circle. Our effort in the application of digital mammography is towards the recognition of circles of *varying radii*. The length of the window should be kept as low as possible. The reason for this is the need for recognition of occluded or distorted circles. For these two cases, the larger the window the less points that contribute to the creation of peaks in the parameter space. For the purpose of this research we chose a window of length equal to twenty as a reasonable compromise between robustness to noise and robustness to missing (occluded) regions.

## 2.6.4 Synthetic Images

### Image Formation

The purpose of using synthetic images is to compare the bisection-based HT with gradient-based HT in terms of robustness and to extract conclusions about the accuracy of the technique as a function of the Signal to Noise Ratio (*SNR*). Our effort was to replicate the image acquisition model presented by Lyvers and Mitchell [93]. According to this model the grey level value of a pixel is given by the following equation

$$I(k, l) = \int_{(k-0.5)\Delta x}^{(k+0.5)\Delta x} \int_{(l-0.5)\Delta y}^{(l+0.5)\Delta y} f(x, y) dx dy, \quad (181)$$

where  $f(x, y)$  denotes a continuous image and  $I(k, l)$  denotes the digital image. An image containing a circle, modeled with this scheme, is presented in Figure 44(a). One can easily notice the smoothness of the edges of the circle. The image shown in this figure is corrupted with Gaussian noise of standard deviation 1. The background value is equal to 120 and the foreground 140. The *SNR* is calculated by the following equation [93]:

$$SNR = 20 \log_{10} \frac{c}{\sigma} \quad (182)$$

where  $c$  is the contrast and  $\sigma$  is the standard deviation of the noise.

### Examples

Figure 44(b) shows the output of edge detection and thresholding of the image of Figure 44(a). The edge detection method of choice was the one proposed by Canny [19]. In Figure 44(c) we present the resulting voting space in a mesh format. The coordinates of the identified peak were (128, 128) and they exactly coincided with the coordinates of the original circle. In Figure 44(d) we show the radius histogram and in Figure 44(e) we show the filtered histogram. Finally, in Figure 44(f), we superimpose the detected circle on the original image.

The same process was followed for an image of low *SNR* and the results are presented in Figure 45. We note here that:

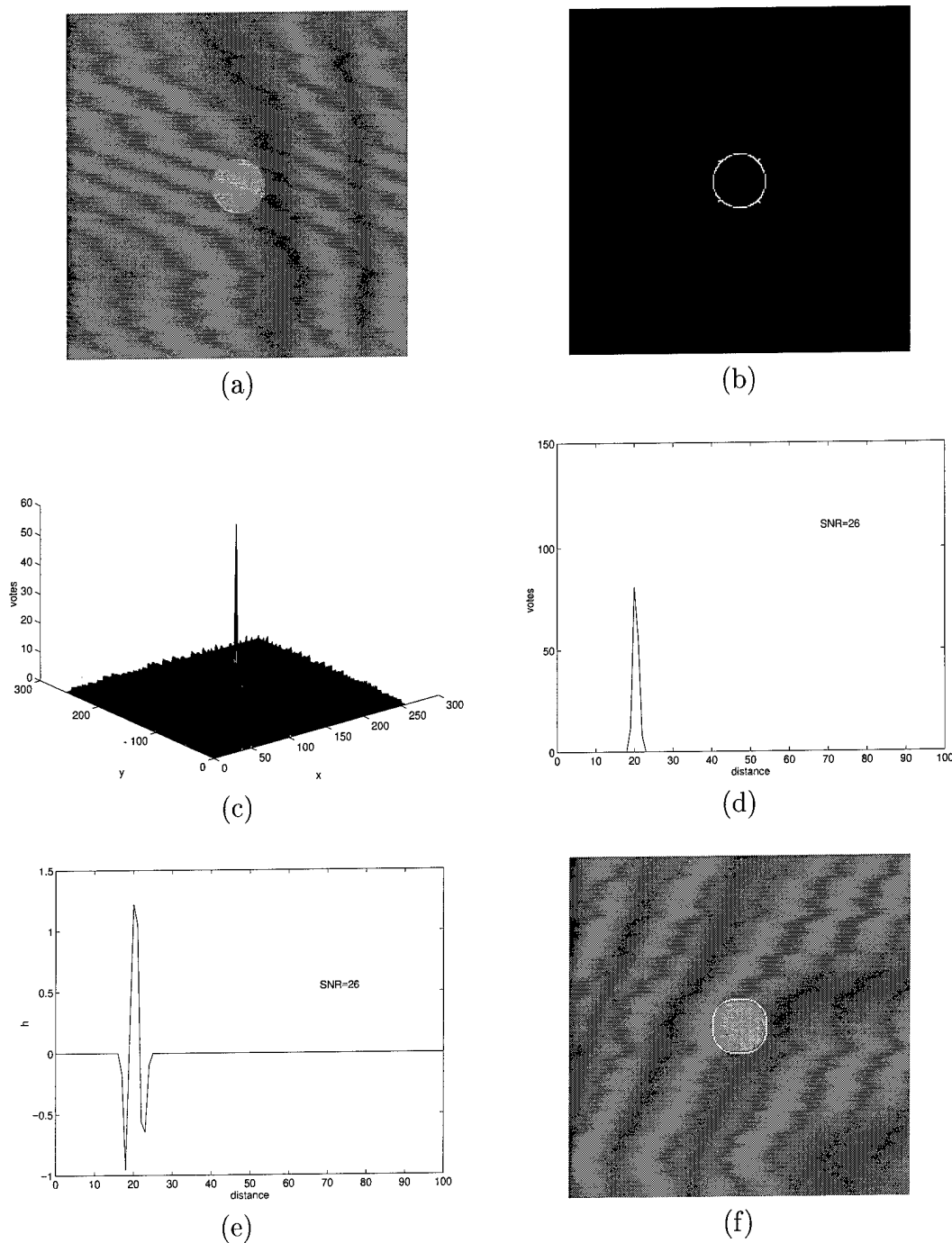


Figure 44: Figure illustrating the key stages of the algorithm for a high  $SNR$  synthetic image ( $SNR = 26$ ). (a) Original image with a circle at its center. The radius of the circle was 20 pixels; (b) edge detected image; (c) plotting of the voting space in a mesh format. The peak indicates the possible existence of a circle. The coordinates of the peak provide an estimation of the coordinates of the center of the circle; (d) histogram of the feature pixels as a function of their distance from the identified center (peak in the voting space); (e) filtered histogram. The identified peak was at  $r = 20$ ; (f) detected circle superimposed on the original image.

1. The number of votes of the center is smaller in the low  $SNR$  case than in the high  $SNR$  case, and
2. The spreading of the votes in the radius histogram is higher in the low  $SNR$  (more bins share the same votes).

To counter these two problems, one has to set the vote threshold of the 2D HT to lower values and to use a larger window during the histogram filtering stage. A final note is the fact that the value of the filtered histogram is greater than one is attributed to the votes due to noisy pixels.

### Comparison of bisection-based with gradient-based HT

Robustness in the context of noise, complex backgrounds and accuracy in the determination of the center of the circle are the two main issues of concern when using the bisection-based 2D Hough transform for center detection. We use the robustness ratio  $P$ , defined by Equation 180, and the distance of the detected center and the true center as measures of the performance of each algorithm. We performed experiments for a wide range of SNRs (2 to 26) for various radii. Experiments were repeated for 20 times and mean values are plotted in the figures that follow.

In Figure 46 we present plots of robustness  $R$ , as defined by Equation 180 as a function of the  $SNR$  for circle of radii 20, 30, and 50 pixels. Comparing the  $R$  of the gradient-based HT and the bisection-based HT, we can conclude that the former is advantageous for high  $SNRs$ , while the latter is far more robust for low  $SNRs$ .

It was also found that the accuracy in the determination of the center for low  $SNR$  with the bisection-based method did not decrease considerably while the gradient-based method gave large errors (see Figure 47).

### RMI phantom

Figure 48(a) shows a radiographic image of an RMI 156 (Gamex Inc., Middleton, WI) mammographic accredited phantom. Three types of mammographic features are visible within this phantom image, simulating difference sized circular malignant masses as well as small microcalcifications and fibers of varying contrast. Small circumscribed lesions are one of the more important signs of breast cancer which may be detected by mammography in asymptomatic women. The detection of lesions in digital mammograms, is one of the main areas digital mammography has focused on [32, 99]. Mammographic phantoms like the one shown in Figure 48(a) are used to measure the performance of mammographic systems and a mammography unit needs to be able to visualize at least four fibers, three groups of

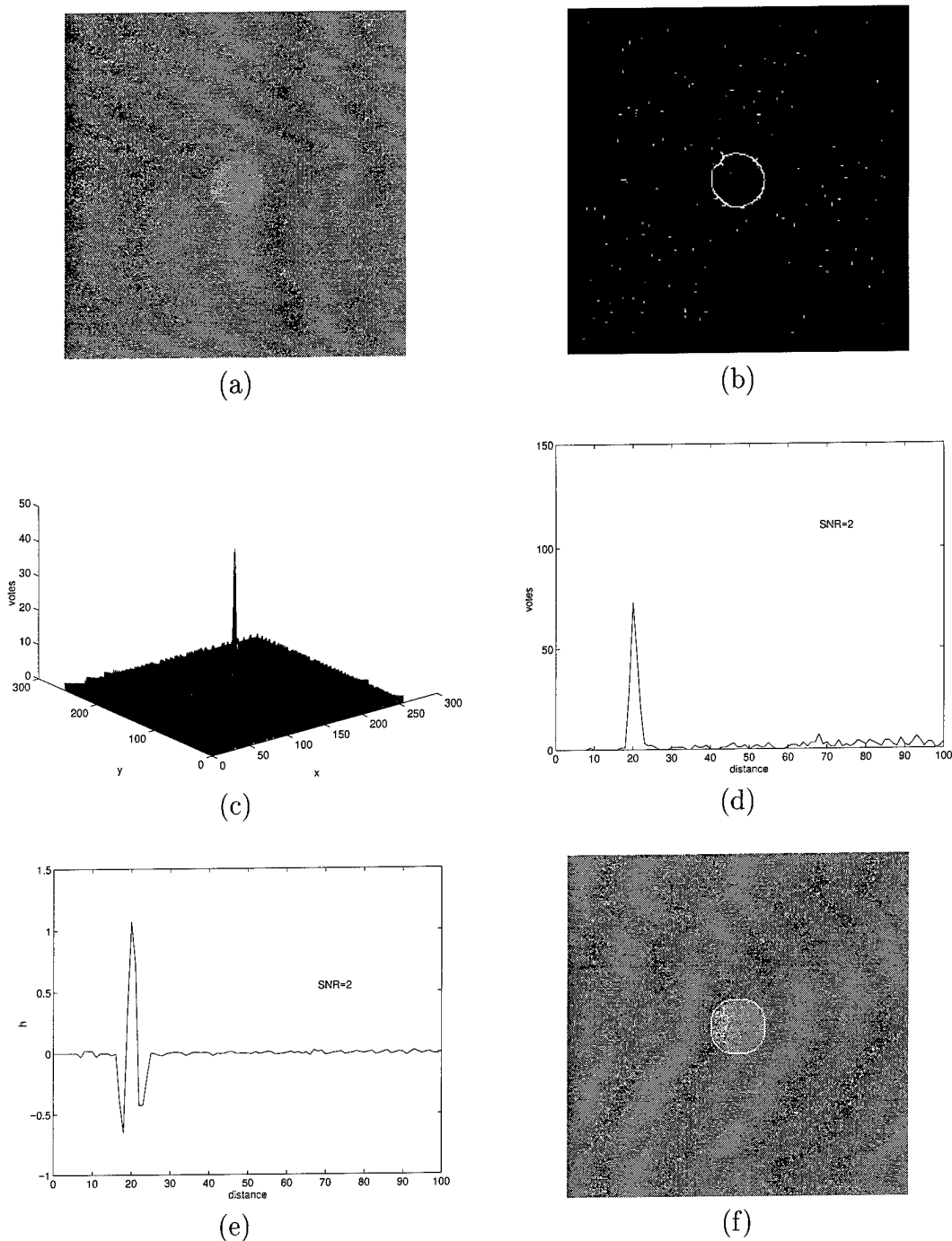


Figure 45: Figure illustrating the key stages of the algorithm for a low  $SNR$  synthetic image ( $SNR = 2$ ). (a) Original image with a circle at its center. The radius of the circle was 20 pixels; (b) edge detected image; (c) plotting of the voting space in a mesh format. The peak indicates the possible existence of a circle. The coordinates of the peak provide an estimation of the coordinates of the center of the circle; (d) histogram of the feature pixels as a function of their distance from the identified center (peak in the voting space); (e) filtered histogram. The identified peak was at  $r = 20$ ; (f) detected circle superimposed on the original image.

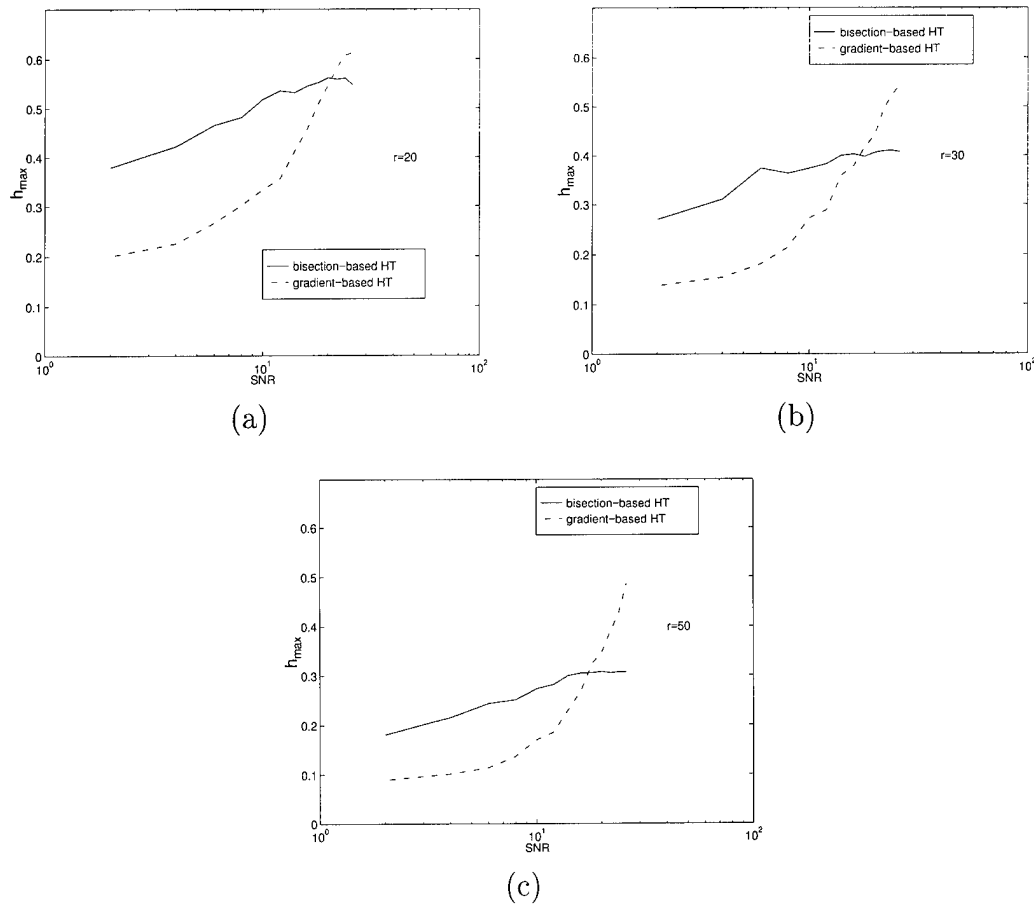


Figure 46: Plots of the mean value of the maximum value of the filtered histogram  $h_{max}$ , for 20 experiments, as a function of the  $SNR$ . Continuous lines denote results using the bisection-based HT, while dashed lines denote results of the gradient-based HT, for circles of radii: (a)  $r = 20$ ; (b)  $r = 30$ ; and (c)  $r = 50$ . While for high  $SNR$  (i.e.  $SNR > 20$ ) the gradient-based HT works better, for low  $SNR$  drop the bisection-based HT gives better results.

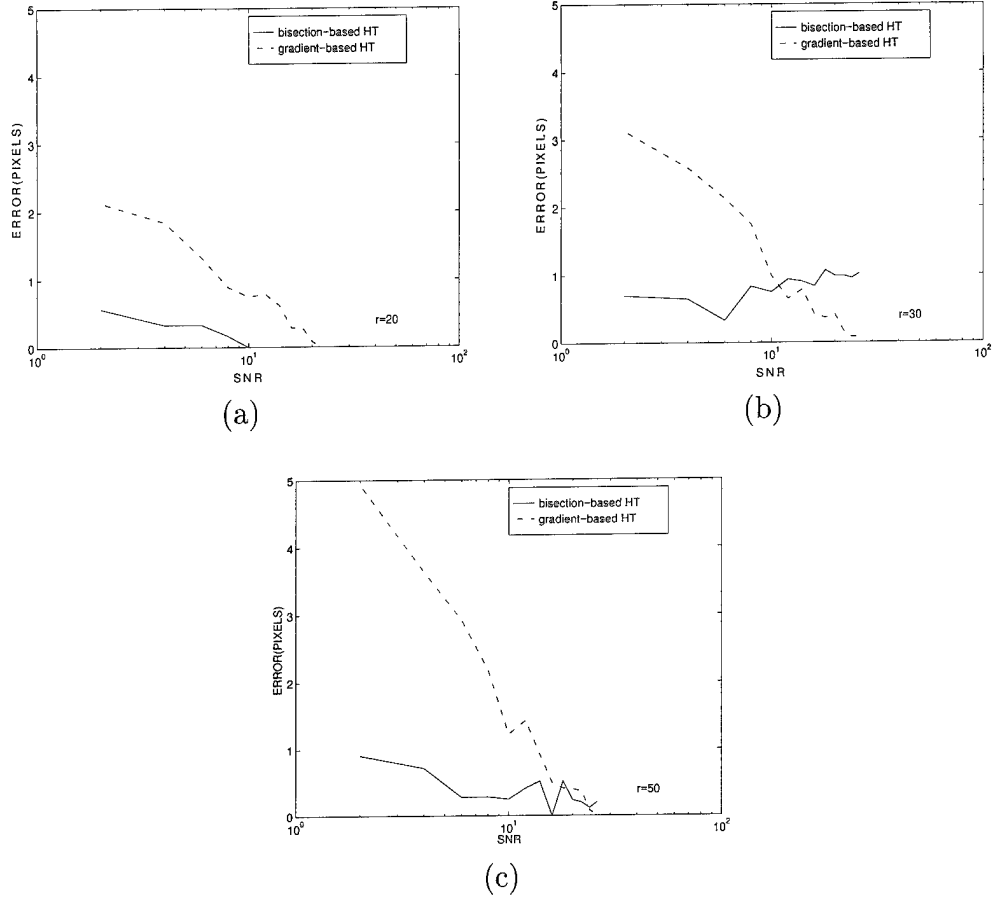


Figure 47: Plots of the error in the determination of the center of the circle with the use of the bisection-based method (continuous lines) and the gradient-based method (dashed lines) for circles of radii: (a)  $r = 20$ ; (b)  $r = 30$ ; and (c)  $r = 50$ . For most of the cases the error of the gradient-based method is larger than the error of the bisection-based method. As shown in these plots, the error of the bisection-based method for a wide range of noise levels remained smaller than a pixel.



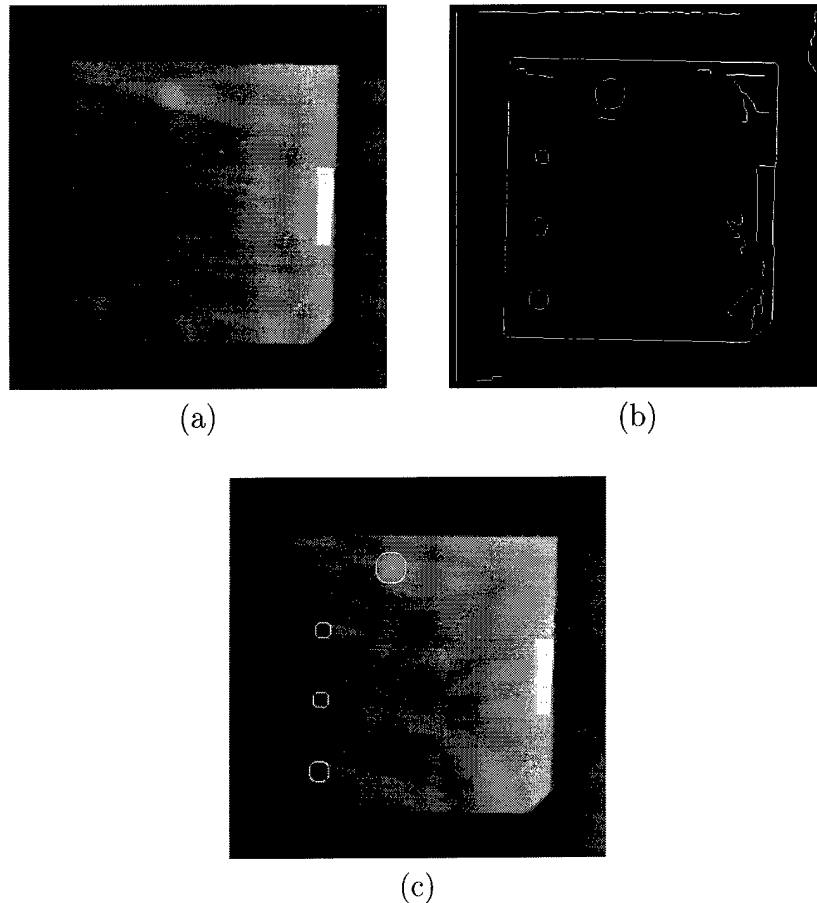


Figure 48: Figure illustrating the steps of the algorithm on a digitized phantom. (a) Image of a phantom; (b) output of Canny edge detection; (c) detected circles superimposed on the original image.

microcalcification specks and three masses to pass the American College of Radiology accreditation program in the United States. The size of the image is 512 by 512 pixels. Figure 48(b) shows the results of edge detection. Eighteen circle centers were detected, when the voting space was searched for local maxima that had more than 30 votes. After radius histogramming, histogram filtering with the filter Equation 179, four circles were detected. These four circles are superimposed on the original image in Figure 48(c).

### Mammogram

Figure 49(a) shows a 512 by 512 image that was cropped from a digital mammogram (mammogram md005.pgm from the MIAS MiniMammography Database-<http://s10d.smb.man.ac.uk/MIASmini.html>). The image contains two overlapping circular masses (biopsy proven) located at the lower middle part of the image. Figure 49(b) shows the output of the Canny edge detection, while Figure 49(c) shows the

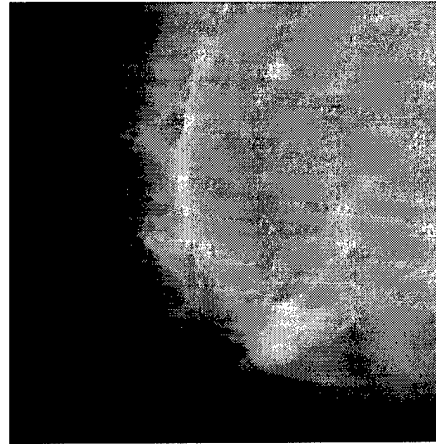
circular masses detected by the proposed algorithm. Both masses were identified but as can be seen from the same figure a false positive also appears (circle located at the top middle of the image).

### 2.6.5 Summary

We presented a two-step HT for the detection and localization of circles in digital images. The first used a bisection-based 2D HT. Experiments with synthetic and real data suggest that the bisection-based HT is more robust to noise than the gradient-based HT while being more accurate. Davies was the first to use this property for the localization of centers of circles [91]. His implementation is very efficient in terms of computation but since it used only vertical and horizontal chords, it lacked robustness needed for analysis in complex backgrounds.

The second step used radius histogramming to detect a circle and extract its radius. Enhancement of local maxima with the use of a Laplacian filter as suggested by Kierkegaard [100], had a significant shortcoming that it did not normalize for the increasing number of pixels belonging to a bin when moving away from a center of the circle (see Equations 178 and 177). Furthermore, there was no intuitive method for the selection of the threshold of the filtered histogram. Our filtering scheme, considers votes due to noisy pixels, shape distortions and normalizes to account for the dependence of the number of pixels on the radius of a circle.

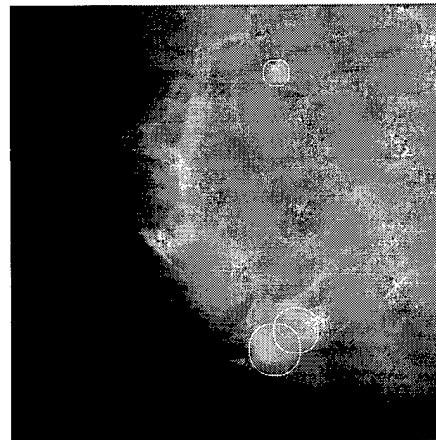
Sample results were provided and showed that the method is capable of detecting circular masses of varying sizes and shape distortions in complex background, including dense breast.



(a)



(b)



(c)

Figure 49: Figure illustrating the steps of the algorithm on a digitized mammogram. (a) Digitized mammogram; (b) output of Canny edge detection; (c) detected circles superimposed on the original image. The algorithm correctly detected the two overlapping masses while giving a false positive (top middle of the image).

### 3 Conclusions

During the past year, we have made significant progress in the development of a methodology for accomplishing adaptive contrast enhancement by multiscale representations. Our studies have demonstrated that features extracted from multiresolution representations can provide a mechanism for the local emphasis of salient and subtle features of importance to mammography. The improved contrast of mammographic features makes these techniques appealing for computed aided diagnosis (CAD) and screening mammography. Screening mammography examinations are certain to grow substantially in the next few years, and analytic methods that can assist general radiologists in reading mammograms shall be of great importance.

In the paragraphs below, we summarize our progress and identify future directions of research to be carried out during the final year of our investigation.

The one-dimensional discrete dyadic wavelet transform was extended to higher-order derivatives and even-order spline functions, and an improved initialization procedure was developed. In comparison to the originally employed initialization [7], our method showed significantly better performance at finer scales of analysis (of importance for mammographic features including microcalcifications).

A multiscale spline derivative-based transform was constructed as an approximation to a steerable dyadic wavelet transform. The transform used  $x$ - $y$  separable filters in a perfect reconstruction filter bank and enabled fast directional analysis of images without the introduction of artifacts due to translation and rotation invariance. Such artifacts are inherent to traditional methods of wavelet analysis.

Preliminary image fusion results showed promise for producing enhanced images with a more familiar appearance to radiologists.

We have described a parallel algorithm to accomplish high-speed interactive multiscale processing for enhancement within ROIs in digital mammograms. The implementation relied upon foundation-level libraries, XIL, on top of a SPARCstation's SX frame buffer. We showed the amount of speedup for the method compared to traditional linear convolution and a conventional FFT approach. Our multiscale approach employing splitting in decomposition and merging in reconstruction was shown to be efficient.

We showed that the continuous wavelet transform of a band-limited signal can be estimated from samples of a signal. If a mother wavelet is compact in the frequency domain, effectively no approximation error was observed experimentally.

We described a modified dyadic wavelet transform, which was able to decompose a signal onto an arbitrary scale and reconstruct perfectly the signal from that scale. We

presented a special class of spline wavelet filters, which allowed us to use a high order scaling function without significantly increasing the computational complexity of analysis.

The importance of arbitrary scale analysis was demonstrated by digital radiographs of a mammography phantom and digitized mammograms. The study showed that the algorithm was able to detect very subtle masses, which were rated to be almost invisible by radiologist specializing in mammography.

An enhancement algorithm relying on multiscale wavelet analysis and extracting oriented information at each scale of analysis was investigated. The evolution of wavelet coefficients across scales characterized the local shape of irregular structures. Using oriented information to detect the features of an image appears to be a promising approach for enhancing complex and subtle structures of the breast. Steerable filters which can be rotated at arbitrary orientations can reliably find visual cues within each spatial-frequency sub-band of an image. "Coherence measure" and "dominant orientation" clearly helped us to discriminate features from complex surrounding tissue typical in mammograms.

Existing and previous multiscale enhancement approaches [72, 32, 59] attempted to enhance an image by detecting edges. Unfortunately, most edge detection algorithms can not distinguish between "authentic" edges and phantom edges. In contrast, this algorithm relied upon a coherence map and phase information which resulted in an enhancement naturally close to the original image. This type of artifact free enhanced image can provide more "obvious" (and familiar) visual cues for radiologists.

In this chapter we have presented the tools that will allow us to compute a multivoice undecimated wavelet transform exactly for an arbitrary wavelet such as the sine-Gabor wavelet. Furthermore, we have shown that the sine-Gabor wavelet is indeed a wavelet frame. However, we found that having nearly optimum time-frequency localization and obtaining a tight frame at the same time is not possible. Both constraints are important because they allow for "good" time-frequency localization and efficient reconstruction, respectively. We have also learned that a better (tighter) frame can be obtained by adding voices. This will not only ease the reconstruction but is also a requirement for applications in time-frequency analysis. Our current work is concerned with extending the sine-Gabor wavelet to the multivoice framework and with the computation of multivoice undecimated wavelet transforms. Our future work includes wavelet reconstruction in the multivoice/undecimated framework as well as extending our results to the analysis of two dimensional data such as images.

Finally, we presented a two-step HT for the detection and localization of circles in digital images. The first used a bisection-based 2D HT. Experiments with synthetic and real data suggest that the bisection-based HT is more robust to noise than the

gradient-based HT while being more accurate. Davies was the first to use this property for the localization of centers of circles [91]. His implementation is very efficient in terms of computation but since it used only vertical and horizontal chords, it lacked robustness needed for analysis in complex backgrounds.

The second step used radius histogramming to detect a circle and extract its radius. Enhancement of local maxima with the use of a Laplacian filter as suggested by Kierkegaard [100], had a significant shortcoming that it did not normalize for the increasing number of pixels belonging to a bin when moving away from a center of the circle (see Equations 178 and 177). Furthermore, there was no intuitive method for the selection of the threshold of the filtered histogram. Our filtering scheme, considers votes due to noisy pixels, shape distortions and normalizes to account for the dependence of the number of pixels on the radius of a circle.

Sample results were provided and showed that the method is capable of detecting circular masses of varying sizes and shape distortions in complex background, including dense breast.

These results are encouraging and continue to support that wavelet based image processing algorithms can play an important role in improving the imaging performance of digital mammography. We believe we are now close to our original schedule and anticipate that we shall complete the objectives described in Phase IV and Phase V this year.

## References

- [1] S. L. Parker, T. Tong, S. Bolden, and P. A. Wingo, "Cancer statistics 1996", *CA-A Cancer Journal for Clinicians*, vol. 46, no. 1, pp. 5-27, 1996.
- [2] R. A. Smith, "Epidemiology of breast cancer", in *A Categorical Course in Physics, Technical Aspects of Breast Imaging*, A. G. Haus and M. J. Yaffe, Eds. The 79th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), 1993, pp. 21-33.
- [3] P. C. Johns and M. J. Yaffe, "X-ray characterization of normal and neoplastic breast tissues", *Physics in Medicine and Biology*, vol. 32, no. 6, pp. 675-695, 1987.
- [4] M. J. Yaffe, R. J. Jennings, R. Fahrig, and T. R. Fewell, "X-ray spectral considerations for mammography", in *A Categorical Course in Physics, Technical Aspects of Breast Imaging*, A. G. Haus and M. J. Yaffe, Eds. The 79th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), 1993, pp. 63-72.
- [5] A. Laine, W. Huda, D. Chen, and J. Harris, "Segmentation of masses using continuous scale representations", in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds., Amsterdam, The Netherlands, 1996, pp. 447-450, Elsevier.
- [6] G. Qu, W. Huda, B. G. Steinbach, and J. C. Honeyman, "An ROC comparison between digital mammography and screen-film using an anthropomorphic breast phantom", *Medical Imaging*, 1997, to appear.
- [7] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710-732, 1992.
- [8] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.
- [9] M. Unser, A. Aldroubi, and S. J. Schiff, "Fast implementation of the continuous wavelet transform with integer scales", *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3519-3523, 1994.
- [10] C. E. Shannon, "Communication in the presence of noise", *Proceedings of the IRE*, vol. 37, no. 1, pp. 10-21, 1949.

- [11] M. Unser, A. Aldroubi, and M. Eden, "Polynomial spline signal approximations: Filter design and asymptotic equivalence with Shannon's sampling theorem", *IEEE Transactions on Information Theory*, vol. 38, no. 1, pp. 95–103, 1992.
- [12] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 277–285, 1991.
- [13] M. Holschneider, R. Kronland-Martinet, J. Morlet, and Ph. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform", in *Wavelets: Time-frequency Methods and Phase Space*, J. M. Combes, A. Grossmann, and Ph. Tchamitchian, Eds., Springer-Verlag, Berlin, Germany, 1990, pp. 286–304.
- [14] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets", *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, 1992.
- [15] C. de Boor, *A Practical Guide to Splines*, Springer-Verlag, New York, NY, 1978.
- [16] J. Fan and A. Laine, "Multiscale contrast enhancement and denoising in digital radiographs", in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds., Boca Raton, FL, 1996, pp. 163–189, CRC Press.
- [17] A. Laine, J. Fan, and S. Schuler, "A framework for contrast enhancement by dyadic wavelet analysis", in *Digital Mammography*, A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns, Eds., Amsterdam, The Netherlands, 1994, pp. 91–100, Elsevier.
- [18] A. Laine, J. Fan, and W. Yang, "Wavelets for contrast enhancement of digital mammography", *IEEE Engineering in Medicine and Biology Society Magazine*, vol. 14, no. 5, pp. 536–550, 1995.
- [19] J. Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [20] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.
- [21] P. Perona, "Deformable kernels for early vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 5, pp. 488–499, 1995.



- [22] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms", *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, 1992.
- [23] Y. Hel-Or and P. C. Teo, "Canonical decomposition of steerable functions", in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 1996, pp. 809–816.
- [24] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson, "Overcomplete steerable pyramid filters and rotation invariance", in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, 1994, pp. 222–228.
- [25] A. Karasarisidis and E. Simoncelli, "A filter design technique for steerable pyramid image transforms", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA, 1996, vol. 4, pp. 2389–2392.
- [26] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation", in *Proceedings of the IEEE International Conference on Image Processing*, Washington, D.C., 1995, vol. 3, pp. 444–447.
- [27] I. Koren, A. Laine, F. Taylor, and M. Lewis, "Interactive wavelet processing and techniques applied to digital mammography", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA, 1996, vol. 3, pp. 1415–1418.
- [28] I. Koren, A. Laine, and F. Taylor, "Image fusion using steerable dyadic wavelet transform", in *Proceedings of the IEEE International Conference on Image Processing*, Washington, D.C., 1995, vol. 3, pp. 232–235.
- [29] A. Laine, I. Koren, W. Yang, and F. Taylor, "A steerable dyadic wavelet transform and interval wavelets for enhancement of digital mammography", in *Wavelet Applications II*, Proceedings of SPIE, Orlando, FL, 1995, vol. 2491, pp. 736–749.
- [30] A. J. Jerri, "The Shannon sampling theorem—its various extensions and applications: A tutorial review", *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.
- [31] I. Koren and A. Laine, "A discrete dyadic wavelet transform for multidimensional feature analysis", in *Time-Frequency and Wavelet Transforms in Biomedical Engineering*, M. Akay, Ed., IEEE Press, New York, NY, 1997.

- [32] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis", *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 725–740, 1994.
- [33] J. K. Aggarwal, *Multisensor Fusion for Computer Vision*, Springer-Verlag, Berlin Heidelberg, Germany, 1993.
- [34] A. Toet, "Image fusion by a ratio of low-pass pyramid", *Pattern Recognition Letters*, vol. 9, no. 4, pp. 245–253, 1989.
- [35] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion", in *Proceedings of the Fourth International Conference on Computer Vision*, Berlin, Germany, 1993, pp. 173–182.
- [36] T. Ranchin, L. Wald, and M. Mangolini, "Efficient data fusion using wavelet transform: the case of spot satellite images", in *Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, Proceedings of SPIE, San Diego, CA, 1993, vol. 2034, pp. 171–178.
- [37] H. Li, B. S. Manjunath, and S. K. Mitra, "Multi-sensor image fusion using the wavelet transform", in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, 1994, vol. 1, pp. 51–55.
- [38] L. J. Chipman, T. M. Orr, and L. N. Graham, "Wavelets and image fusion", in *Wavelet Applications in Signal and Image Processing III*, Proceedings of SPIE, San Diego, CA, 1995, vol. 2569, pp. 208–219.
- [39] Sun Microsystems, Inc., *SPARCstation 10SX Graphics Technology*, 1993.
- [40] Sun Microsystems, Inc., *XIL Programmer's Guide*, 1994.
- [41] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [42] D. Marr and E. Hildreth, "Theory of edge detection", *Proceedings of the Royal Society of London*, vol. 207, no. 1167, pp. 187–217, 1980.
- [43] M. Vrhel, C. Lee, and M. Unser, "Fast computation of the continuous wavelet transform through oblique projections", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA, 1996, vol. 3, pp. 1459–1462.

- [44] M. Vrhel, C. Lee, and M. Unser, "Fast continuous wavelet transform", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, MI, 1995, vol. 2, pp. 1165–1168.
- [45] A. Aldroubi, M. Unser, and M. Eden, "Cardinal spline filters: Stability and convergence to the ideal sinc interpolator", *Signal Processing*, vol. 28, no. 2, pp. 127–138, 1992.
- [46] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape", *SIAM Journal on Mathematical Analysis*, vol. 15, no. 4, pp. 723–736, 1984.
- [47] A. Aldroubi and M. Unser, "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory", *Numerical Functional Analysis and Optimization*, vol. 15, no. 1–2, pp. 1–21, 1994.
- [48] G. Strang, "Wavelets and dilation equations: A brief introduction", *SIAM Review*, vol. 31, no. 4, pp. 614–627, 1989.
- [49] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions", *Communications in Mathematical Physics*, vol. 110, no. 4, pp. 601–615, 1987.
- [50] C.K. Chui and J. Z. Wang, "On compactly supported spline wavelets and a duality principle", *Transactions of the American Mathematical Society*, vol. 330, no. 2, pp. 903–915, 1992.
- [51] M. Unser, P. Thévenaz, and A. Aldroubi, "Shift-orthogonal wavelet bases using splines", *IEEE Signal Processing Letters*, vol. 3, no. 3, pp. 85–88, 1996.
- [52] A. G. Haus and M. J. Yaffe, *A categorical course in physics, Technical aspects of breast imaging*, The 80th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), Oak Brook, IL, 1994.
- [53] J. Lim, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1990.
- [54] R. L. Egan, *Breast Imaging*, W. B. Saunders Company, Philadelphia, PA, 1988.
- [55] R. McLelland, "Stellate lesions of the breast", *Recent Results in Cancer Research*, vol. 119, pp. 24–27, 1990.

- [56] W. P. Kegelmeyer Jr., J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions", *Radiology*, vol. 191, no. 2, pp. 331–337, 1994.
- [57] Y. Xing, W. Huda, A. Laine, J. Fan, B. Steinbach, and J. Honeyman, "Comparison of a dyadic wavelet image enhancement algorithm with unsharp masking and median filtering", in *Medical Imaging: Image Processing*, Proceedings of SPIE, San Diego, CA, 1995, vol. 2434, pp. 718–729.
- [58] W. Qian, L. P. Clarke, M. Kallergi, H.-D. Li, R. Velthuisen, R. A. Clark, and M. L. Silbiger, "Tree-structured nonlinear filter and wavelet transform for microcalcification segmentation in mammography", in *Biomedical Image Processing and Biomedical Visualization*, Proceedings of SPIE, San Jose, CA, 1993, vol. 1905, pp. 509–520.
- [59] W. B. Richardson Jr., "Nonlinear filtering and multiscale texture discrimination for mammograms", in *Mathematical Methods in Medical Imaging*, Proceedings of SPIE, San Diego, CA, 1992, vol. 1768, pp. 293–305.
- [60] L. Li, W. Qian, and L. P. Clark, "X-ray medical image processing using directional wavelet transform", in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, GA, 1996, vol. 4, pp. 2251–2254.
- [61] M. Kass and A. Witkin, "Analyzing oriented patterns", *Computer Vision, Graphics, and Image Processing*, vol. 37, no. 3, pp. 362–385, 1987.
- [62] A. R. Rao and B. G. Schunck, "Computing oriented texture fields", in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, 1989, pp. 61–68.
- [63] K. R. Castleman, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, 1996.
- [64] H. G. Barrow and J. M. Tenenbaum, "Recovering intrinsic scene characteristics from images", in *Computer Vision Systems*, A. Hanson and E. Riseman, Eds., New York, NY, 1978, pp. 3–26, Academic Press.
- [65] D. Marr, *Vision*, Freeman, San Francisco, CA, 1982.
- [66] M. J. Shensa, "The discrete wavelet transform: Wedding the à trous and Mallat algorithms", *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2464–2482, 1992.

- [67] A. Laine and C.-M. Chang, "De-noising via wavelet transforms using steerable filters", in *IEEE International Symposium on Circuits and Systems*, Seattle, WA, 1995, vol. 3, pp. 1956–1959.
- [68] J. J. Koenderink and A. J. van Doorn, "Generic neighborhood operators", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, pp. 597–605, 1992.
- [69] N. Karssemeijer and G. M. te Brake, "Detection of stellate distortions in mammograms", *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, pp. 611–619, 1996.
- [70] M. C. Morrone and R. A. Owens, "Feature detection from local energy", *Pattern Recognition Letters*, vol. 6, no. 5, pp. 303–313, 1987.
- [71] R. Owens, S. Venkatesh, and J. Ross, "Edge detection is a projection", *Pattern Recognition Letters*, vol. 9, no. 4, pp. 233–244, 1989.
- [72] A. Laine and S. Song, "Multiscale wavelet representations for mammographic feature analysis", in *Mathematical Methods in Medical Imaging*, Proceedings of SPIE, San Diego, CA, 1992, vol. 1768, pp. 306–316.
- [73] D. Gabor, "Theory of communication", *Journal of the Institution of Electrical Engineers*, vol. 93, no. 26, pp. 429–457, 1946.
- [74] C. K. Chui, *An Introduction to Wavelets*, Academic Press, Boston, MA, 1992.
- [75] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 2091–2110, 1989.
- [76] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.
- [77] O. Rioul and M. Vetterli, "Wavelets and signal processing", *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14–38, 1991.
- [78] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and understanding continuous wavelet transforms", in *Wavelets: Time-frequency Methods and Phase Space*, J. M. Combes, A. Grossmann, and Ph. Tchamitchian, Eds., Springer-Verlag, Berlin, Germany, 1990, pp. 2–20.

- [79] T. S. Lee, "Image representation using 2D Gabor wavelets", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959–971, 1996.
- [80] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design", *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2207–2232, 1992.
- [81] I. Daubechies, "Orthonormal bases of compactly supported wavelets", *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, 1988.
- [82] S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases of  $L^2(\mathbb{R})$ ", *Transactions of the American Mathematical Society*, vol. 315, no. 1, pp. 69–87, 1989.
- [83] V. F. Leavers, "SURVEY: Which Hough transform?", *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 58, pp. 250–264, 1993.
- [84] J. Illingworth and J. Kittler, "SURVEY: A survey of the Hough transform", *Computer Vision, Graphics and Image Processing*, vol. 44, no. 1, pp. 87–116, 1988.
- [85] A. Rosenfeld, *Picture Processing by Computer*, Academic Press, New York, NY, 1969.
- [86] R. O. Duda and P. E. Hart, "Use of the Hough Transformation to detect lines and curves in pictures", *Communications of the Association of Computing Machinery*, vol. 15, no. 1, pp. 11–15, 1972.
- [87] E. R. Davies, "A modified Hough scheme for general circle location", *Pattern Recognition Letters*, vol. 7, no. 1, pp. 37–43, 1988.
- [88] R. K. K. Yip, P. K. S. Tam, and D. N. K. Leung, "Modification of Hough Transform for circles and ellipses detection using a 2-dimensional array", *Pattern Recognition*, vol. 25, no. 9, pp. 1007–1022, 1992.
- [89] D. C. W. Pao, H. F. Li, and R. Jayakumar, "Shapes recognition using the straight line Hough transform: Theory and generalization", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1076–1089, 1992.
- [90] G. L. Foresti, C. S. Regazzoni, and G. Vernazza, "Circular arc extraction by direct clustering in a 3D Hough parameter space", *Signal Processing*, vol. 41, pp. 203–224, 1995.
- [91] E. R. Davies, "A high speed algorithm for circular object location", *Pattern Recognition Letters*, vol. 6, no. 12, pp. 323–333, 1987.

- [92] E. R. Davies, "The effect of noise on edge orientation computations", *Pattern Recognition Letters*, vol. 6, pp. 315–322, 1987.
- [93] V. P. Lyvers and R. Mitchell, "Precision edge contrast and orientation estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 927–937, 1988.
- [94] H. K. Yuen, J. Princen, J. Illingworth, and J. Kittler, "Comparative study of Hough Transform methods for circle finding", *Image and Vision Computing*, vol. 8, no. 1, pp. 71–77, 1990.
- [95] T. M. van Veen and F. C. Groen, "Discretization errors in the Hough transform", *Pattern Recognition*, vol. 14, no. 1–6, pp. 137–145, 1981.
- [96] M. Worring and W. M. Smeulders, "Digitized circular arcs: Characterization and parameter estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 6, pp. 587–598, 1995.
- [97] Z. Kulpa, "On the properties of discrete circles, rings, and disks", *Computer Graphics and Image Processing*, vol. 10, no. 4, pp. 348–365, 1979.
- [98] I. Amir, "Algorithm for finding the center of circular fiducials", *Computer Vision, Graphics, and Image Processing*, vol. 49, no. 3, pp. 398–406, 1990.
- [99] B. R. Groshong and W. P. Kegelmeyer Jr., "Evaluation of a Hough Transform method for circumscribed lesion detection", in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Smith, Eds., Amsterdam, The Netherlands, 1996, pp. 361–364, Elsevier.
- [100] P. Kierkegaard, "A method for detection of circular arcs based on the Hough transform", *Machine Vision and Applications*, vol. 5, pp. 249–263, 1992.
- [101] G. Qu, W. Huda, A. F. Laine, B. G. Steinbach, and J. C. Honeyman, "Can digital mammography improve low-contrast lesion detection?", in *Digital Mammography '96*, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Eds., Amsterdam, The Netherlands, 1996, pp. 451–454, Elsevier.

# A Appendix

## A.1 Initial Condition

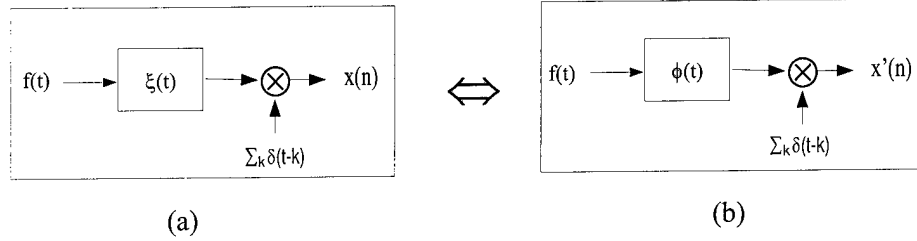


Figure 50: Modify initial condition

A non-ideal acquisition device is shown in Fig 50 (a). The impulse response of the acquisition device is  $\xi(t)$ , where  $\xi(t)$  is a lowpass filter. The approximation of  $\xi(t)$  in  $V_\phi$  space is

$$\xi(t) \approx \tilde{\xi}(t) = \sum_k z(k) \phi(t - k), \quad (183)$$

where  $z(k)$  is the projection of  $\xi(t)$  onto the scaling space. The discrete signal  $x(k)$  is

$$\begin{aligned} x(n) &= \int_{-\infty}^{\infty} f(t) \xi(n - t) dt \\ &\approx \int_{-\infty}^{\infty} f(t) \left( \sum_k z(k) \phi(n - t - k) \right) dt \\ &= \sum_k z(k) \left( \int_{-\infty}^{\infty} f(t) \phi(n - t - k) dt \right) \\ &= \sum_k z(k) x'(k), \end{aligned} \quad (184)$$

where  $x'(n) = \int_{-\infty}^{\infty} f(t) \phi(n - t) dt$ . Given  $x(k)$ ,

$$x'(n) = \sum_k x(k) z^{-1}(n - k). \quad (185)$$

When  $\xi(t)$  is an ideal lowpass filter, it can be approximated as a cardinal spline filter [45]. If the scaling function is a spline of order  $n - 1$ ,

$$\xi(t) \approx \sum_k (b^n)^{-1}(k) \text{beta}(t - k). \quad (186)$$

Therefore,  $z^{-1}(k) = b^n(k)$ ,  $z(k) = (b^n)^{-1}(k)$ . Notice that Equation (185) is similar to projecting a signal onto a scale space as described in Unser's algorithm [43] [44], although the starting point is quite different in our case.



## A.2 Proof of Theorem 3 in Section 2.3.7

In this appendix, we will show that if  $\phi(t) = \beta^{n-1}(t)$ ,  $\phi_1(t) = \beta^{m-1}(t)$ ,

$$P_a(\omega) = \left( \frac{\Psi_a(\omega)}{\Phi(\omega)} \text{Rect}(\omega) \right) * 2\pi\delta_{2\pi}(\omega) \quad (187)$$

as  $n + m \rightarrow +\infty$ .

**Proof:** Note a central B spline is symmetric,  $\beta^{m-1}(t) = \beta^{m-1}(-t)$ . In the oblique projection method,

$$q_a(k) = \int_{-\infty}^{\infty} \psi_a(t) \beta^{m-1}(t-k) dt = \int_{-\infty}^{\infty} \psi_a(t) \beta^{m-1}(k-t) dt = \psi_a * \beta^{m-1} \quad (188)$$

Substituting Equation (188) into Equation (85),

$$\begin{aligned} p_a(k) &= \sum_{i \in \mathcal{Z}} \left( \int_{-\infty}^{\infty} \psi_a(t) \beta^{m-1}(i-t) dt \right) (b^{n+m-1})^{-1}(k-i) \\ &= \int_{-\infty}^{\infty} \sum_{i \in \mathcal{Z}} \psi_a(t) \beta^{m-1}(i-t) (b^{n+m-1})^{-1}(k-i) dt. \end{aligned} \quad (189)$$

Since  $\beta^{m-1} * \beta^{n-1} = \beta^{n+m-1}$ , in the frequency domain, we write

$$\hat{\beta}^{m-1}(\omega) = \hat{\beta}^{n+m-1}(\omega) \frac{1}{\hat{\beta}^{n-1}(\omega)}. \quad (190)$$

Define  $\hat{\beta}^{n-1}(\omega) = \frac{1}{\hat{\beta}^{n-1}(\omega)}$ . Then

$$\begin{aligned} \beta^{m-1}(i-t) &= \beta^{n+m-1} * \dot{\beta}^{n-1}(i-t) \\ &= \int_{-\infty}^{\infty} \beta^{n+m-1}(v) \dot{\beta}^{n-1}(i-t-v) dv \\ \text{let } u &= i-v \\ &= \int_{-\infty}^{\infty} \beta^{n+m-1}(i-u) \dot{\beta}^{n-1}(u-t) du. \end{aligned} \quad (191)$$

Substituting Equation (191) into Equation (189),

$$\begin{aligned} p_a(k) &= \int_{-\infty}^{\infty} \sum_{i \in \mathcal{Z}} \psi_a(t) \left( \int_{-\infty}^{\infty} \beta^{n+m-1}(i-u) \dot{\beta}^{n-1}(u-t) du \right) (b^{n+m-1})^{-1}(k-i) dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{i \in \mathcal{Z}} \psi_a(t) \beta^{n+m-1}(i-u) \dot{\beta}^{n-1}(u-t) (b^{n+m-1})^{-1}(k-i) dt du \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \psi_a(t) \dot{\beta}^{n-1}(u-t) dt \right) \left( \sum_{i \in \mathcal{Z}} \beta^{n+m-1}(i-u) (b^{n+m-1})^{-1}(k-i) \right) du \\ \text{Let } i' &= k-i, \text{ and} \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \psi_a(t) \dot{\beta}^{n-1}(u-t) dt \right) \left( \sum_{i' \in \mathcal{Z}} \beta^{n+m-1}(k-i'-u) (b^{n+m-1})^{-1}(i') \right) du \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} \psi_a(t) \dot{\beta}^{n-1}(u-t) dt \right) \eta^{n+m-1}(k-u) du. \end{aligned} \quad (192)$$

Where  $\eta^{n+m-1}(t) = \sum_{i \in \mathbb{Z}} \beta^{n+m-1}(t-i)(b^{n+m-1})^{-1}(i)$  and  $\eta^{n+m-1}(t)$  is a cardinal spline filter. As  $n+m \rightarrow +\infty$ , it converges to the ideal sinc interpolator. Therefore,

$$P_a(\omega) = \left( \frac{\Psi_a(\omega)}{\Phi(\omega)} \text{Rect}(\omega) \right) * 2\pi\delta_{2\pi}(\omega). \quad (193)$$

■

### A.3 Derivation of Equation 179

In this appendix we will derive the filter shown in Equation 179. We shall assume that we are dealing with an image corrupted with uniform noise and that we are looking for a perfect digital circle. As we noted (see Equation 176) earlier, for  $\Delta r = h$  the maximum number of bins a circle spreads its votes is equal to 2. Therefore, if we pass the radius histogram with the filter

$$t = [ \ 1 \ 1 \ ]$$

we will obtain an estimate of the pixels the circle contains. We would like the filter to be symmetric around the central bin and therefore we choose

$$t = [ \ 1 \ 1 \ 1 \ ].$$

The resulting array contains the number of pixels of the circle plus votes due to noise, other objects etc. We need to enhance our scheme with a module that estimates the votes due to noise and subtracts them. For this purpose we use two more bins (the outer ones). Therefore our filter becomes

$$t = [ \ \alpha \ 1 \ 1 \ 1 \ \beta \ ]. \quad (194)$$

To get values for  $\alpha$  and  $\beta$  we will require that our filter give an unbiased estimator with minimum uncertainty (standard deviation). We assume that noise is spatially uniform. To have an unbiased estimator means that if the image is corrupted with noise and no circular object is present, the resulting (processed) histogram will have mean value equal to zero.

It is easy to see that the  $r$ -th bin of the histogram represents a ring [97] whose radius is equal to  $r$ . Kulpa provided experimental evidence that a good approximation to the number of pixels belonging to such a ring is  $2\pi r$ . This estimation is based on empirical evidence [97] and until now no theoretical justification to it had been found. We verified, through simulations, that this formula holds with adequate accuracy for radii in the range

[h,512h]. It is easy to show, that the number of votes the  $r$ -th bin takes due to noise follows a binomial distribution with mean value:

$$\mu = p(2\pi r) \quad (195)$$

and standard deviation:

$$\sigma = \sqrt{p(1-p)(2\pi r)}, \quad (196)$$

where  $p$  is the probability that a pixel is noisy.

From Equation 195 and the constraint we imposed, that our estimator should be unbiased, we formulate

$$\mu_F = \alpha(r - 2h) + (r - h) + (r) + (r + h) + \beta(r + 2h) = 0. \quad (197)$$

The second equation that we use comes from the requirement of minimization of the standard deviation of the resulting filtered histogram. The standard deviation of the filtered space is given by the following formula:

$$\sigma_F = 2\pi \sqrt{\alpha^2(r - 2h)^2 + (r - h)^2 + (r)^2 + (r + h)^2 + \beta^2(r + 2h)^2}. \quad (198)$$

If we make the following substitutions:

$$X1 = \alpha(r - 2h)$$

$$X2 = r - h$$

$$X3 = r$$

$$X4 = r + h$$

$$X5 = \beta(r + 2h)$$

to the Equations 197 and 198 we obtain

$$X1 + X2 + X3 + X4 + X5 = 0, \quad (199)$$

and

$$\sigma_F = 2\pi \sqrt{X1^2 + X2^2 + X3^2 + X4^2 + X5^2}. \quad (200)$$

If we solve Equation 199 for  $X1$ , and substitute into Equation 200, and differentiate Equation 200 with respect to  $\beta$  we obtain

$$\frac{\partial \sigma_F}{\partial \beta} = (2X5 + 2(X2 + X3 + X4 + X5))f = 0, \quad (201)$$

where  $f$  is a factor that cannot become zero (a constant divided by a function of  $X5$ ).

From Equations 197 and 201 we can compute  $\alpha$  and  $\beta$ . The resulting filter was described in the main part of Section 2.6 (see Equation 179).